

human	Motif 0	
tez1		AKFLHLMMSVYVVELLSFFYVTTTFQKNR
EST2		ISEIEWLVLGKRSNAKMCLSDFEKRKQIFAEEIYWLNSFIIPILQSEFFYITESDDLNR
p123		LKDFRWLFISD--IWFTKHNFENLNQLAICFISWLFRLIPKIIQTFFYCTEISSVT- TREISWMQVET-SAKHFYYFDHEN-IYVLWKLRLWIFEDLVVSLIRCFFVYVTEQQKSYSK * *** **
human	Motif 1	
tez1		LFFYKSVWSKLQSIGIRQHLKRVQLRDVSEAEVRQHREARPALLSRLRFIPKP--DGL
EST2		TVYFRKDIWKLICRPFI-TSMKMEAFEKINENNVRMDTQK-TTLPFAVIRLLPKK--NTF
p123		IVYFRHDTWNKLI TPFIVEYFKTYLVENNVCRNHSYTLS--NFNHSKMRIIPKKSNEF TYYRKNIDVIMKMSI-ADLKKETLAEVQEKEVEEWKKS-LGFAPGKLRLLIPKK--TTF ... * . . * . * . *
human	Motif 2	
tez1		RPVNMDDYVVGARTFRREKRAERLTSRVKALF-SVLNYERA
EST2		RLITN-LRKRFLIKGSKNKKMLVSTNQTLRPVASILKHLINSESSGIPFNLEVYMKLLTF
p123		RIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSTFKIYPTQIADRIKEF RPIMTFNKKIIVNSDRKTTKLTNTKLNLNHLMLKTLKN-RMFKDPFGFAVFNDDVMKKY * * * . *
tez1	Motif 3 (A)	
EST2		KKDLLKHRMFGK-KKYFVRIDIKSCYDRIKQDLMFRIVKK-KLKDPEFVIRKYATIHATS
p123		KQRLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN EEFVCKWKVGQPKLFFATMDIEKCYDSVNREKLSLTKTLKLLSSDFWIMTAQILKRKN * . * * * . * . *

FIG. 1

+

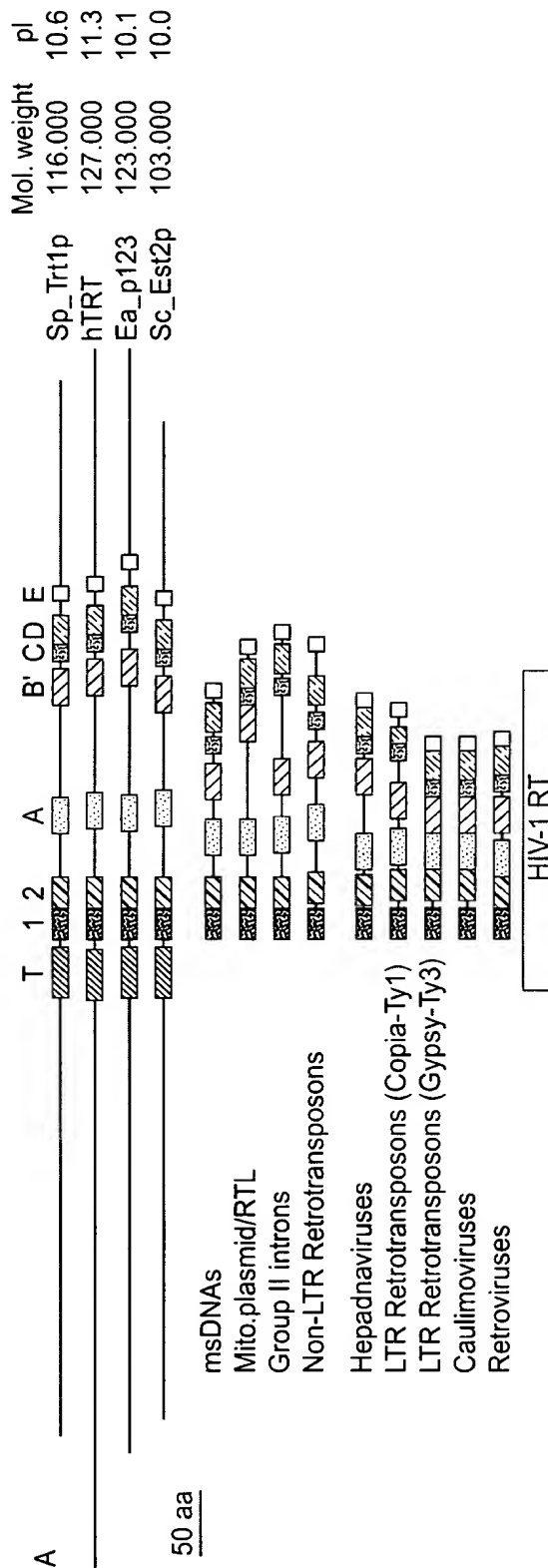


FIG. 2

+

+

3/34

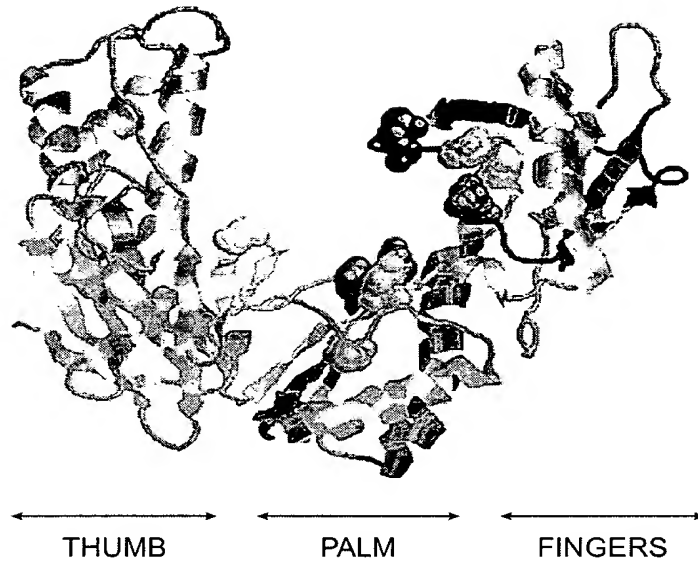


FIG. 3

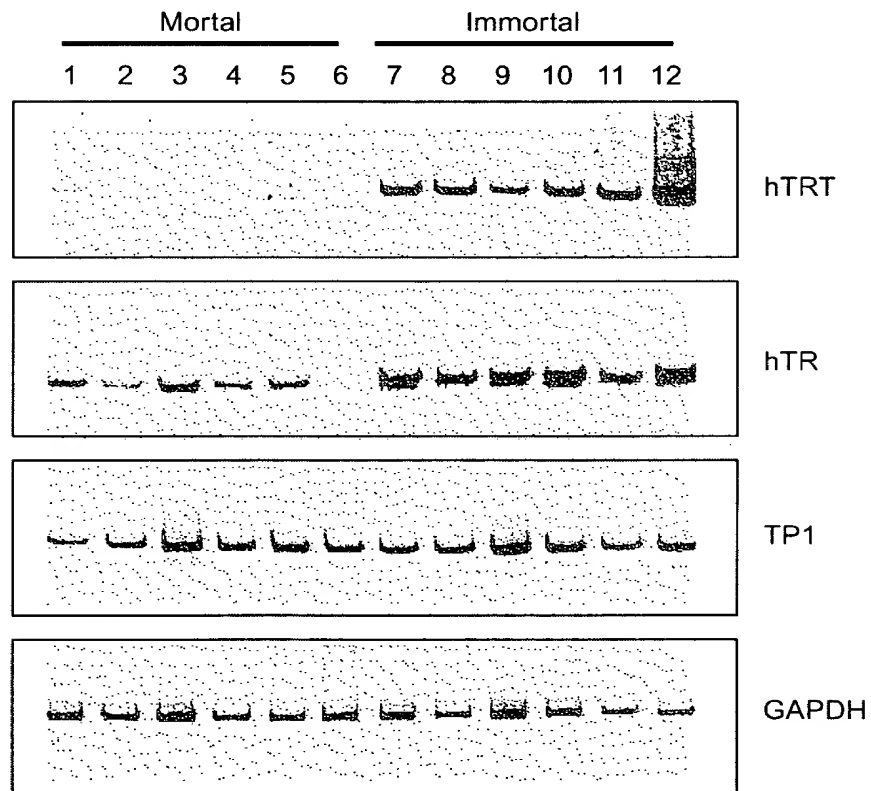


FIG. 5

+

+

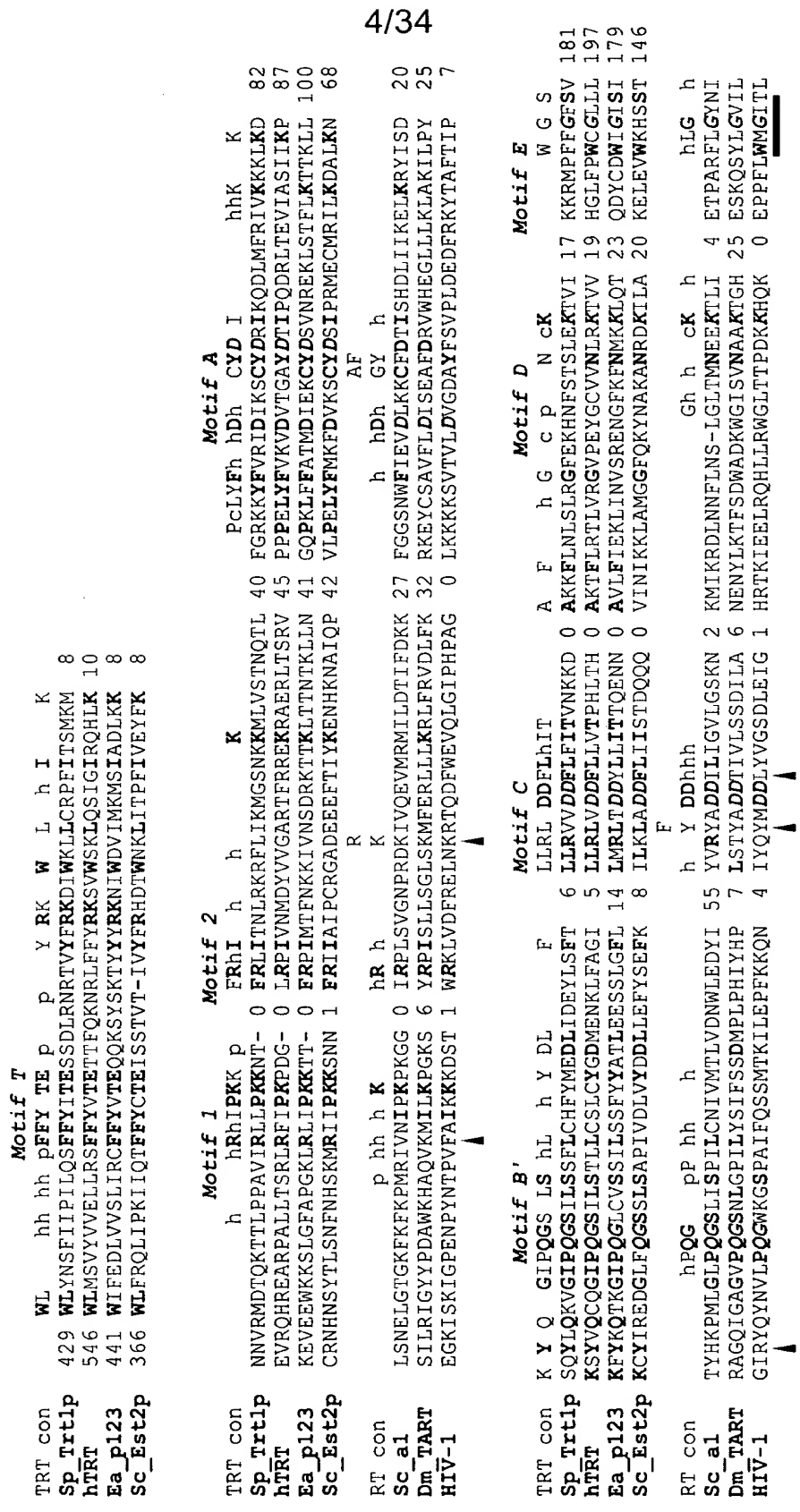
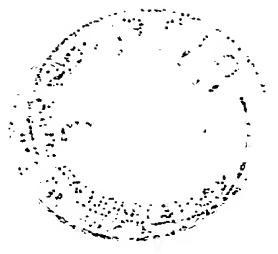


FIG. 4



+

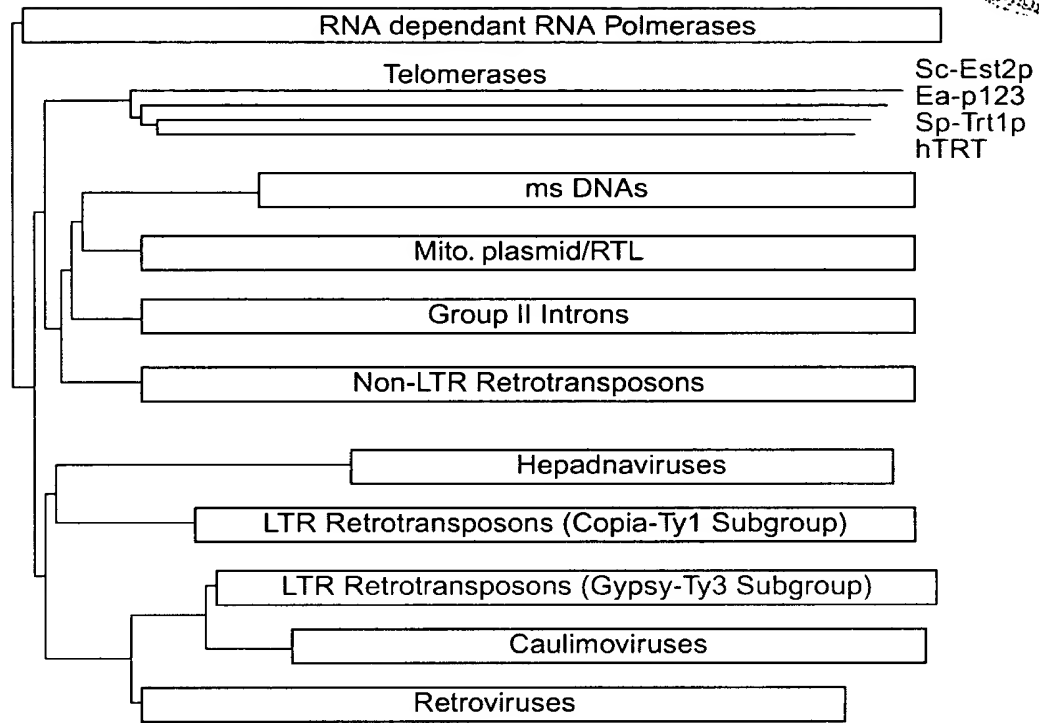


FIG. 6

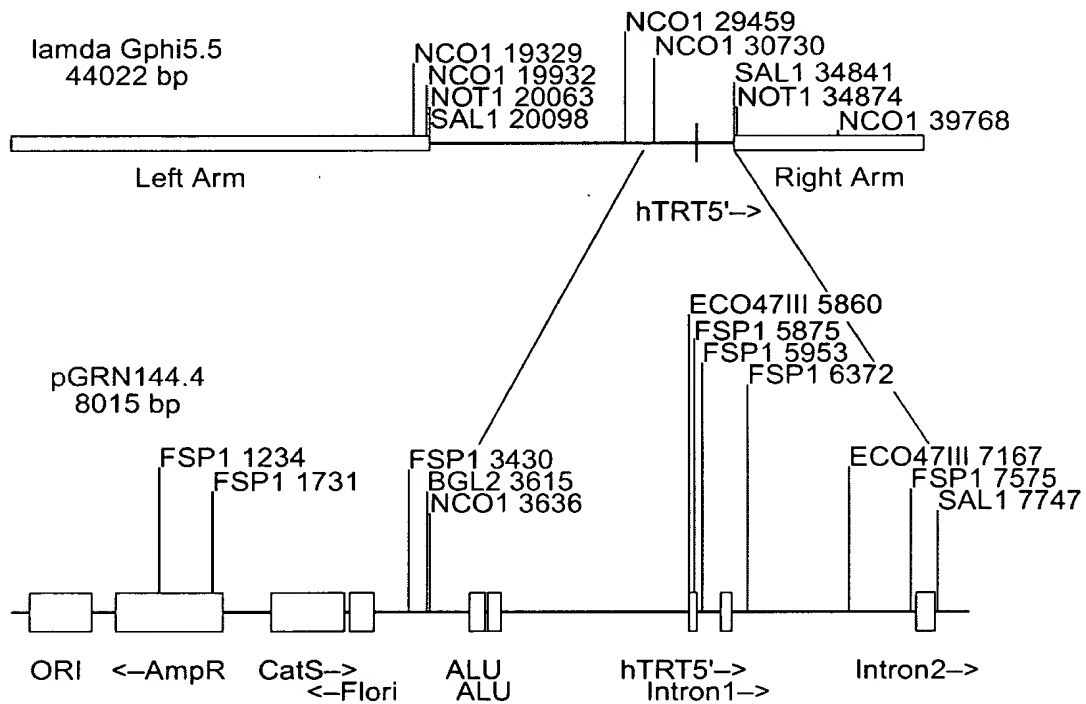


FIG. 7

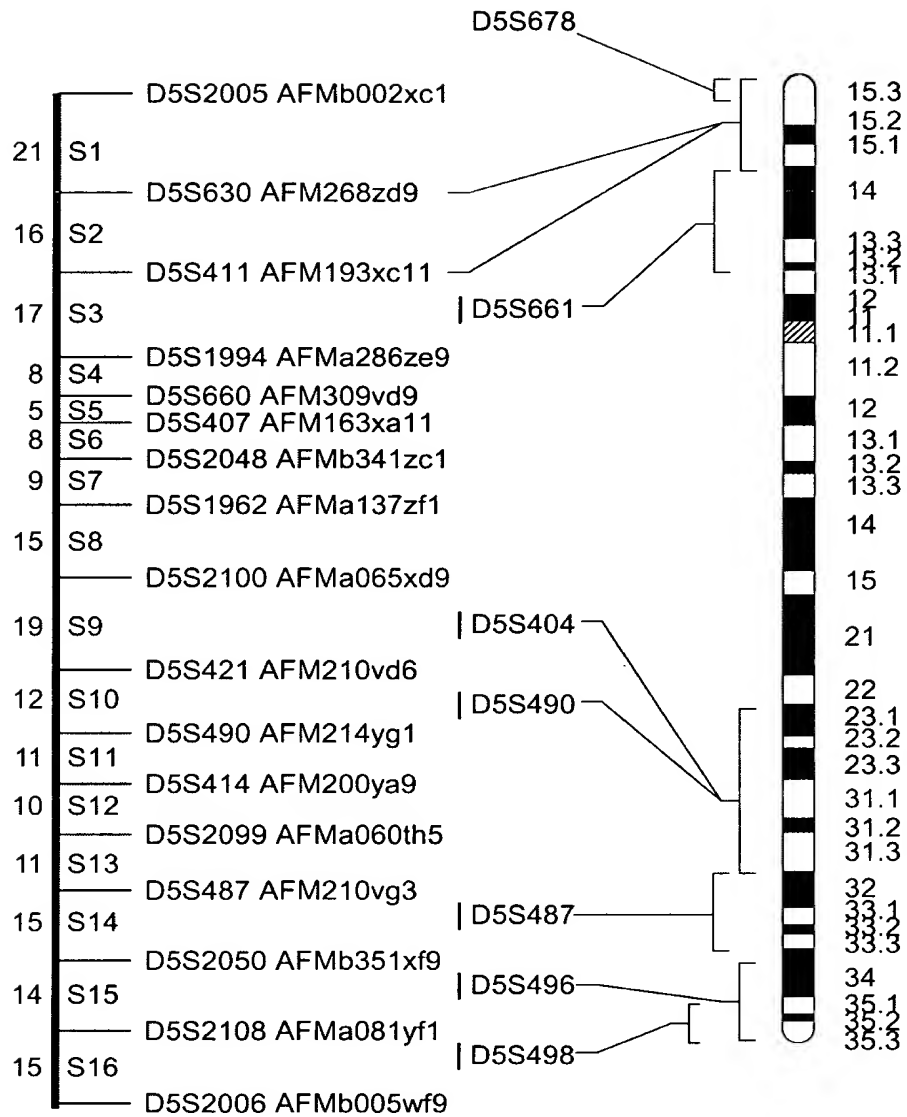


FIG. 8

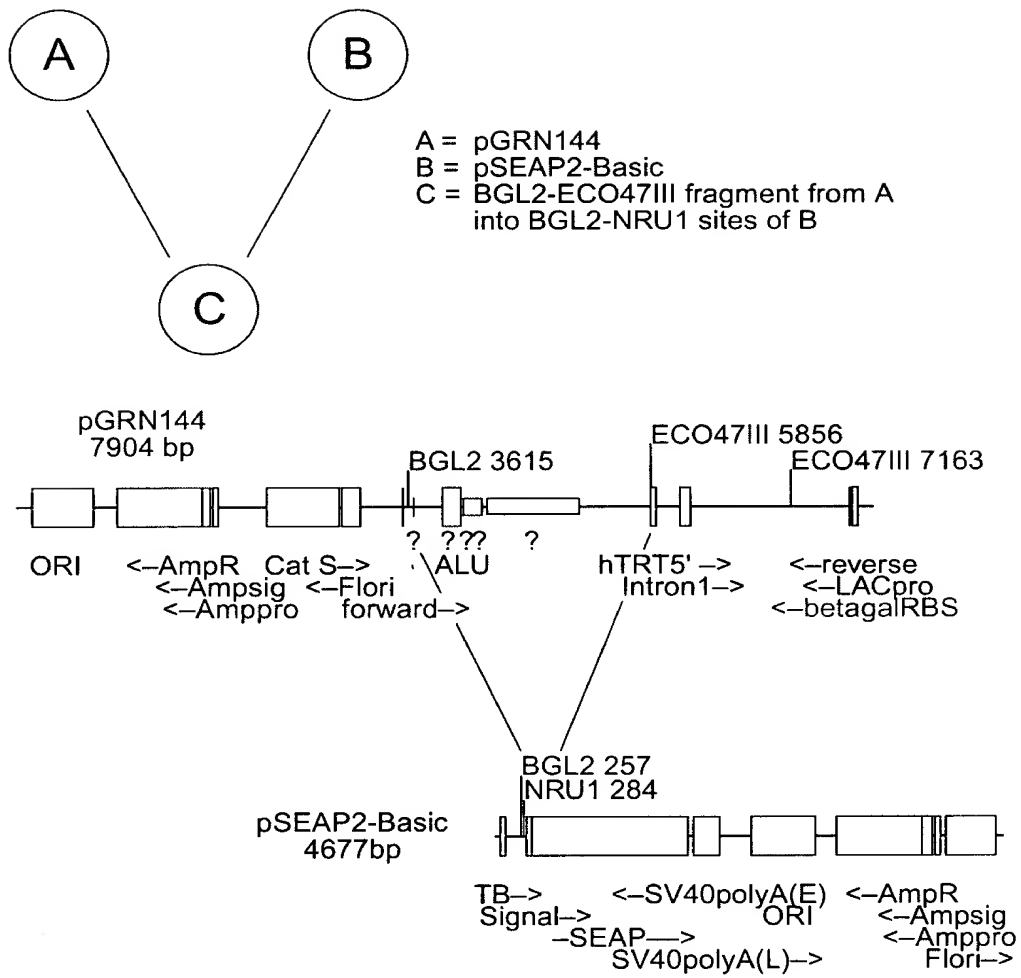


FIG. 9

+

8/34

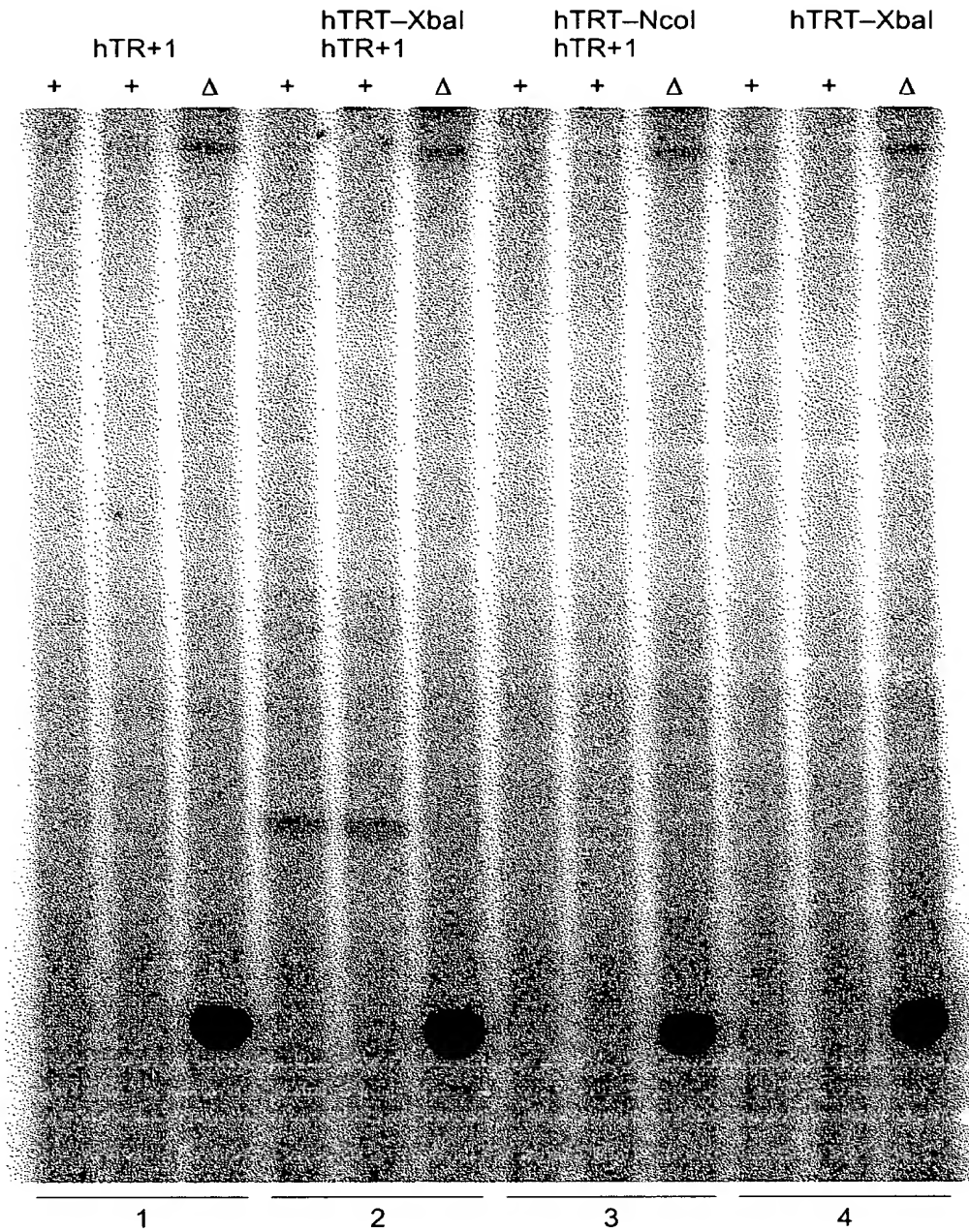


FIG. 10A

+

+

9/34

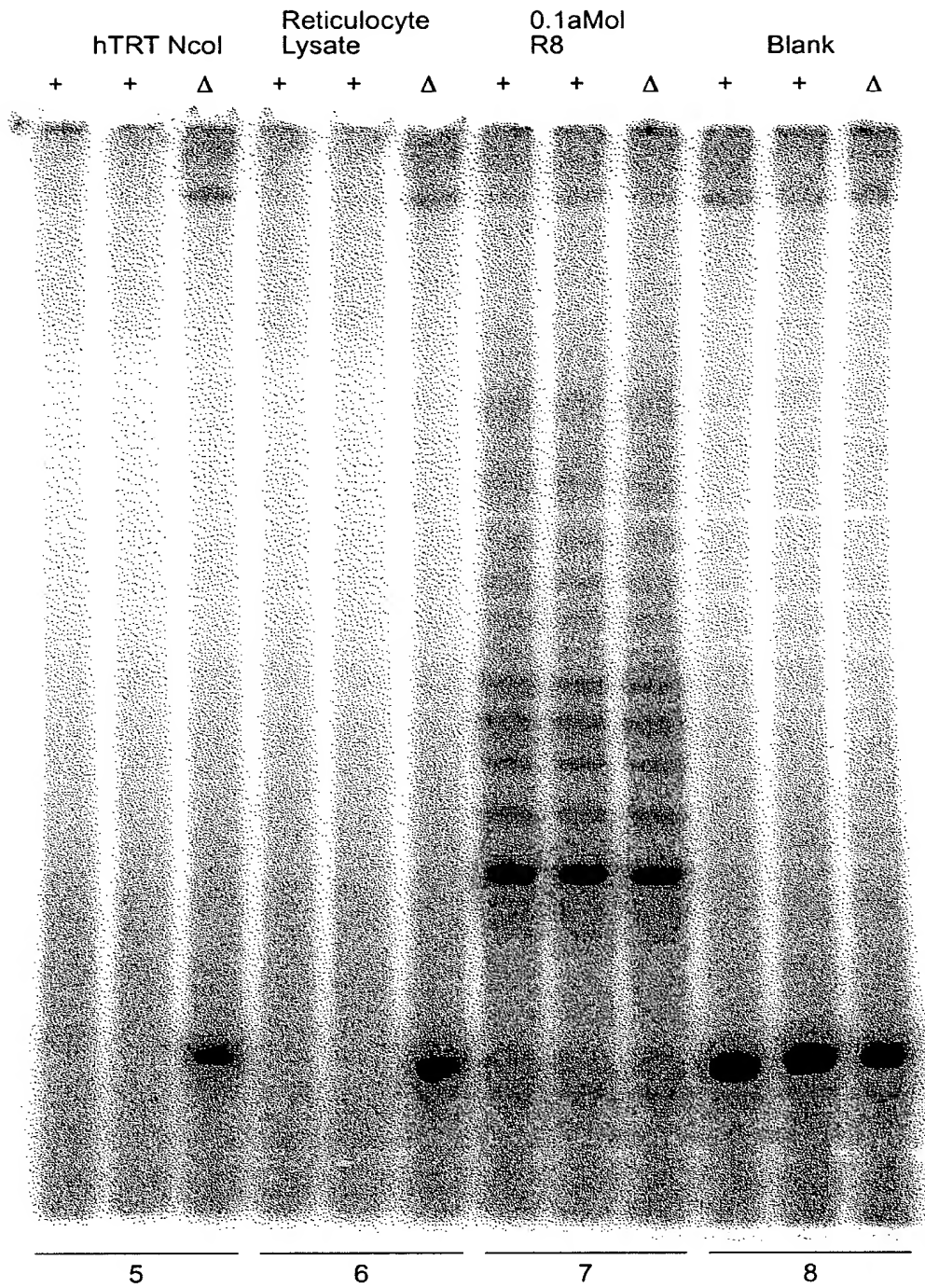


FIG. 10B

+

+

10/34

Telomerase Specific Motifs

	MOTIF T	MOTIF T'
TRT con	Wl	
hTRT	546 WLMSVVVVVLLRSFFVVTETTFQNRLLFFYRKSVWSKLQSIGI 13 EAEVR	E V
spTRT	429 WLYNSFIIPILQSFYITESSDLNRNTVYFRKDIWKLLCRPFI 12 ENNVR	
Ea_p123	441 WIFEDLVVSLIRCFYVTEQOKSYKTYRKNINWDVIMKMSI 12 EKEVE	
Sc_Est2	366 WLFRLIPKIIQTFYCYCTEISSTVT.IVYFRHDTWNKLTTPFI 9 ENNV	

Telomerase RT Motifs (Fingers)

	MOTIF 1	MOTIF 2	MOTIF A	MOTIF B'
TRT con	R ipKk	fr I	p lyF D cyd i	Y q GipQGs lS l Y
hTRT	11 SRLRFIPKPDG 0 LRPIV	69 PELYFVKVDVTGAYDTI 104 YVQCGIPQGSILSTLLCSLCY		
spTRT	10 AVIRLLPKKNT 0 FRLIT	66 RKKYFVRIDIKSCYDRI 99 YLQKVGIPQGSILSSFLCHFYM		
Ea_p123	10 GKRLIPKKT 0 FRPIM	67 PKLFFATMDIEKCYDSV 117 YKQTKGIPQGLCVSSILSSFYI		
Sc_Est2	13 SKMRIIPKKS 2 FRIIA	68 PELYFMKFDVKSCYDSI 85 YIREDEGLFQGSLSAPIVDLIV		
RT con	p hh h k	hr h	h hDh AF h GY	hpQG pp hh h

Telomerase RT Motifs (Palm, Primer Grip)

	MOTIF C	MOTIF D	MOTIF E
TRT con	lllrl Ddfl it	g	w g s l
hTRT	15 LLLRLVDDFLlLV 15 GVPEYGCVVNLKRTVV 24 WCGLLLDTRTL 192		
spTRT	16 VLLRVVDDFLFIT 15 GFEKHNFSTSLKTVI 22 FFGFSVNMRL 176		
Ea_p123	24 LLMRLTDDYLLIT 15 VSRENGFKFNMKLIQT 28 WIGISIDMKTL 174		
Sc_Est2	18 LILKLADDFLIIS 15 GFQKYNAKANRDKILA 25 WKHSSTWNNFH 141		
RT con	h Y DDhhh	Gh h ck h hLG h	

FIG. 11



+

+

11/34



181 GGACCCGGCGGCTTTCCGCGCGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGC
CCTGGGCCGCGCGAAAGGCGCGCGACCACCGGGTCACGGACCACACGCACGGGACCCTGCG

NFkB_CS1
GGGRQTYYQC
NFkB-MHC-I.2
TGGGCTTCCCC

241 ACGGCCGCCCCCGCCGCCCCCTCCTTCCGCCAGGTGGGCCTCCCCGGGGTCCGGCGTCCG
TGCCGGCGGGGGGCGGCGGGGGAGGAAGGCGGTCCACCCGAGGGGGCCCCAGCCGCAGGC

Intron1

301 GCTGGGGTTGAGGGCGGCCGGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTC
CGACCCCAACTCCCGCCGGCCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

NFkB_CS1
GGGRQTYYQC
NFkB_CS2
RGGGRMTYYCC
Topo_II_cleavage_site
RNYNNCNGYNGKTNYN

*****>

361 AGGGCGCTTCCCCCGCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGG
TCCCGCGAAGGGGGCGTCCACAGGACGGACTTCTCGACCACCGGGCTCACGACGTCTCC

FIG. 12

+

```

1   AAAACCCCAA AACCCCAAAA CCCCTTTT TAG AGCCCTGCAG TTGGAAATAT
51  AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAAACCTAA
101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
251 TTAGAAGATA TTAAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG
301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTCTT
401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTCA
451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA
501 CTCAAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
551 ATTGGAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
601 GCGAACTTCT GAAGGAACTC TTGTTCAATT TTGCGGGAAT AACGTTTTTG
651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
801 ATATGAAATC AAGAACCAGA ATATTTTATT GCACTCATTT TAATAGAAAT
851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTT
901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
1001 CTTGAGAAAG TCAAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG
1051 TCCTCTTCCA GAAAATTGGC GGAACGGAA ACAAAAAATC GAAAACCTGA
1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGC GCTGTTTAGC
1151 TACACAACTG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTCTA
1201 CAATATACTC CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTC
1251 AAAAGAAAGT TAAGAAATAT GTGGAATAA ACAAGCATGA ACTCATTCAC
1301 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
1351 GGTTGAGACC TCTGCAAAGC ATTTTATTA TTTTGATCAC GAAAACATCT
1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
1451 CTGATTAGAT GATTTTCTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
1601 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT
1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
1701 TAAATTCAGA CCGGAAGACT ACAAATTAAC CTACAAATAC GAAGTTATTG
1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
1801 TTTTGGATTG GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
1851 AGTTTGTGTT CAAATGGAAG CAAGTTGGAC AACCAGAACT CTTCTTTGCA
1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAATATC
1951 AACATTCCTA AAAACTACTA AATTACTTTC TTCAGATTTT TGGATTATGA
2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC
2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
2101 TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG
2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
2201 AGAAATTAAT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTTG
2251 CCAATATAAT TACATTAAC TTAATGGGAA GTTTTATAAA CAAACAAAAG
2301 GAATTCCTCA AGGTCTTTGA GTTTCATCAA TTTTGTGCAT ATTTTATTAT
2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

```

FIG. 13A

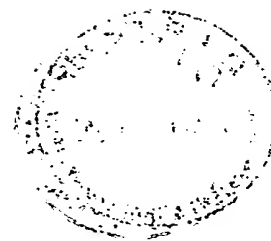
2401 CCCTGAAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
 2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
 2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
 2551 GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
 2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
 2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
 2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
 2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
 2801 CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACCTCT
 2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
 2901 AATACAAGGA CCACTTTAAG AAGAACCTAG CTATGAGCAG TATGATCGAC
 2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
 3001 CCTTGTGTGC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG
 3051 ACTTTTTCCT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA
 3101 AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA
 3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
 3201 CTATTCTAAC TTATTTTGGG AAGTTAATTT TCAATTTTTG TCTTATATAC
 3251 TGGGGTTTTG GGGTTTTGGG GTTTTGGGG

FIG. 13B

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD
 51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
 101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
 151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGG
 201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNMKSRTTR IFYCTHFNRN
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAM
 301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHელი
 401 KNLLLEKINT REISWMQVET SAKHFYFDH ENIYVLWKL RWIFEDLVVS
 451 LIRCFYVTE QQKSYSKTY YRKNIWVIM KMSIADLKKE TLAEVQEKEV
 501 EEWKKS LGFA PGKLRLIPKK TFRPIMTFN KKIVNSDRKT TKLTNTNKL
 551 NSHLMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
 651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ
 701 RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFYY
 751 ATLEESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIKEL
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
 851 SIDMKTALM PNINLRIEGI LCTLNLMQTT KKASMWLKKK LKSFLMNNIT
 901 HYFRKTTITE DFANKTLNKL FISGGYKYM CAKEYKDHF KNLAMSSMID
 951 LEVSKI IYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
 1001 KYIFNRVCM LKAKEAKLKS DQCQSLIQYD A

FIG. 14

+



```

1  ggtagcagatttacttcttcttctcctaagctaattggttctcctcgaacgctcctaaactctctggaataatttttacaaga 80
81  actcaataaacaataccaagtcaaatccaataatgaagtggttattagtagcgatcgataatttctatttttactcggtcgta 160
161  ccaagtataaggacaaaagaacaacttcttccccctaaagactttactttattattacttttcaaatatatttcg 240
241  ggttcgcttacttttaactcggtgtagtctttagctgtagctttagccaacccggtgttttctaccccgcatatggatat 320
321  agctcttgtagtagtcacagagaatccttacaaactctctgatgagactatattagattacagtcogtgcataatc 400
401  ttaacatggagccttacacatttagatgagtcacgtgcagatgaggtatttggatcatccaacgtttgaccttgaaaaag 480
481  gttgataattatttgcaaaatctctcactacaactccttaacgcggttttattttctattttctattctcatgtgtt 560
561  atbgagatatcaaaaatttctatccactacaactccttaacgcggttttattttctattttctattctcatgtgtt 640
641  ccaaatatgtagtctcgtattaggttttccggttttactcctggaatcgtaaccttttccactattccccctaata 720
721  ataactctaaattagttcgtttataattgatagtagtagaagaattggtgattctactcgtgtaattgttattagtttaa 800
801  gatactttgcaaaacatttattagctatcatatataaaaaaaatccataattataaataattcaatatttgcgggtc 880
881  actatttataaaacggttatgatcagtaggacactttgcatatatatagttatgcttaattggttacttgaacttgc 958

959  ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
1  M   T   E   H   H   T   P   K   S   R   I   L   R   F   L   E   N   Q   Y   V   20

1019  TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
21  Y   L   C   T   L   N   D   Y   V   Q   L   V   L   R   G   S   P   A   S   S   40

1079  TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
41  Y   S   N   I   C   E   R   L   R   S   D   V   Q   T   S   F   S   I   F   L   60

1139  CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
61  H   S   T   V   V   G   F   D   S   K   P   D   E   G   V   Q   F   S   S   P   80

1199  AAA TGC TCA CAG TCA GAG gtatatatatattttgttttctatttctattcgggtagctaataatgggcag 1272
81  K   C   S   Q   S   E   86

1273  CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA 1332
87  L   I   A   N   V   V   K   Q   M   F   D   E   S   F   E   R   R   N   L   106

1333  CTG ATG AAA GGG TTT TCC ATG gtaaggtaattcctaattgtgaaatatttaccctgcaattactgtttcaaaaga 1405
107  L   M   K   G   F   S   M   113

1406  ttgtatttaaccgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT 1469
114  N   H   E   D   F   R   A   M   H   V   N   G   V   Q   N   128

```

FIG. 15A

+

+



3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT 3960
 799 R M P F F G G F S V N M R S L D T L L A C 818

 3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG 4020
 819 P K I D E A L F N S T S V E L T K H M G 838

 4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtaactgaataataatagctgacaaataatcag A TCG 4089
 839 K S F F Y K I L R S 848

 4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT 4149
 849 S L A S F A Q V F I D I T H N S K F N S 868

 4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA 4209
 869 C C N I Y R L G Y S M C M R A Q A Y L K 888

 4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagtacttattttaactaga 4274
 889 R M K D I F I P Q R M F I T D 903

 4275 aaagtcattaattaaccccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC 4339
 904 L L N V I G R K I W K K L A 917

 4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401
 918 E I L G Y T S R R F L S S A E V K W 935

 4402 ggtctcgagacttcagcaatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA 4468
 936 L F C L G M R D G L K 946

 4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT 4528
 947 P S F K Y H P C F E Q L I Y Q F Q S L T 966

 4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA ATA 4588
 967 D L I K P L R P V L R Q V L F L H R I 986

 4589 GCT GAT TAA tgtcatttttcaatttattatatatacatccctttattactgggtgtctttaacaataattattactaagtata 4665
 987 A D * 989

FIG. 15E

+

+



4666 gctgacccccaaagcatactataggatttctagtaaagtataaaataatctcgttattagtttttgattgacttgctt 4745
4746 ttatccttatacttttaagaagattgacagtgggttgctgactactgcccacatgcccatataacgggagtggttaaca 4825
4826 ttaaaagtaatacatgaggttaactctccttcatcttagaataaggaaagtggtttctataatgaataatgcccgacta 4905
4906 atgcaaaaagacgaagattatcttaacaaggggatttaagcatatccgaaggaaaaagagagtaatatccccagtggt 4985
4986 gttgaagaaaagcaaggataatttgaacaagctctgcagatgacaggctaaatttggtagccgaatttggtaaaagc 5065
5066 ccagggttaccatggtggccggccttgctactgagacgaaagaaactaaggatagtttgaaataactaatagctcattta 5145
5146 atgtcttataaagggttttctcctgacttcaatttgcattgggtgaaagaaatagtttaagccattattggat 5225
5226 tccgaaatagccaaatttctggltccctcaaagcggaaagtctaaagaacttatgaagcttatgaggttcaaaaactcc 5305
5306 tcctgatttaaggaggaaatctccaccgatgaggaaatggatagcttatcagctgctgaggagagccctaatttttgc 5385
5386 aaaaaagaaaatatcatgggagacatctcttgatgaatcagatcgggagagtatcccagcggtatccttgatgtaata 5465
5466 acttctatttctgaaatgtatgggtcctactgtcgtctcgtagctctacgcagttaaagtaccaaagggtacc 5544

FIG. 15F

+

1 gcagcgctgc gtccctgctgc gcacgtggga agccctggcc ccggccaccc ccgcgatgcc
 61 gcgcgctccc cgctgccgag cctgtgcgctc cctgtgcgctc agccactacc gcgaggtgct
 121 gccgctggcc acgttcgtgc ggccgctggg gcccagggc tggcggctgg tcgagcgagg
 181 ggacccggcg gctttccgcg cgctggtggc ccagtgcctg gtgtgcgtgc cctgggacgc
 241 acggccgccc cccgcccggc cctccttccg ccaggtgtcc tgctgaagg agctggtggc
 301 ccgagtgtgt cagaggctgt ggcagcgcg ggcgaagaac gtgctggcct tcggcttcgc
 361 gctgctggac ggggcccgcg ggggcccccc cgaggccttc accaccagcg tgcgcagcta
 421 cctgcccaac acggtgaccg acgcactgcg ggggagcggg gcgtgggggc tgctgctgcg
 481 ccgctgtggc gacgacgtgc tggttcacct gctggcacgc tgccgctctc ttgtgctggt
 541 ggctcccagc tgcgcctacc aggtgtcgcg cccgcccgtg taccagctcg gcctgccac
 601 tcaggccccg cccccgccac acgctagtgg accccgaagg cgtctgggat gcgaacgggc
 661 ctggaaccat agcgtcaggg aggcgggggt cccctggggc ctgccagccc cgggtgcgag
 721 gaggcgcggg ggagtgcca gccgaagctt gccgttgccc aagaggccca ggctggcgcg
 781 tgcccctgag ccggagcgga cgcccttggt cgagggggtc tgggcccacc tggcaggac
 841 gcgtggaccg agtgaccgtg gtttctgtgt ggtgtcacct gccagacccg ccgaagaagc
 901 cactcttttg gagggtgcgc tctctggcac gcgccactcc caccatccg tgggcccga
 961 ccaccacgcg ggccccccat ccacatcgcg gccaccagct ccctgggaca ggcctgttcc
 1021 cccggtgtac gccgagacca agcacttccct ctactcctca ggcgacaagg agcagctgcg
 1081 gccctccttc ctactcagct ctctgaggcc cagcctgact ggcgctcgga ggctcgtgga
 1141 gaccatcttt ctgggttcca ggcctggat tgccggccct gccagggact ccccgcaggt tccccgcct
 1201 gccccagcgc tactggcaaa tgcggccctt gtttctggag ctgcttggga agcacgcgca
 1261 gtgcccctac ggggtgctcc tcaagacgca ctgcccgtg cgagctgcgg tcaccccagc
 1321 agccggtgtc tgtgcccggg agaagcccca gggctctgtg gcggcccccg aggaggagga
 1381 cacagaccgc cgtgcctgg tgcagctgct cccgagcac agcagccctc ggcaggtgta
 1441 cggcttcgtg cgggcctgct tgcggcggtc ggtgccccca ggctctggg gctccaggca
 1501 caacgaacgc cgcttccctc ggaacaccaa gaagtctatc tccctgggga agcatgccaa
 1561 gctctcgctg caggagctga cgtggaagat gagcgtgcgg gactgcgctt ggctgcgag
 1621 ggaagcagag gttggctgtg ttccggccgc agagcacctg ctgcgtgagg agatcctggc
 1681 caagtctctg cactggctga tgagtgtgta cgtcgtcgag ctgctcaggt ctttctttta
 1741 tgtcacggag accacgtttc aaaagaacag gctctttttc taccggaaga gtgtctggag
 1801 caagtgtcaa agcattgga tccagacgca ctgaaagag gtgcagctgc gggagctgtc
 1861 ggaagcagag gtcaggcagc atcgggaaag caggcccgcg ctgctgacgt ccagactccg
 1921 cttcatcccc aagcctgacg ggctgcggcc gattgtgaac atggactacg tcgtgggagc
 1981 cagaagtttc cgcagagaaa agaggcgcc gcgtctcacc tccaggggtg aggcactgtt
 2041 cagcgtgctc aactacgagc gggcgccgag ccccgccctc ctgggctggg ctgtgtggg
 2101 cctggacgat atccacaggg cctggcgcac ctctgtgctg cgtgtgcggg cctgtgcggg
 2161 gccgcctgag ctgtactttg tcaaggtgga tgtgacgggc gcgtacgaca cccaggaccc
 2221 ggacaggctc acggaggtca tcgccagcat catcaaacc cagaacacgt actgcgtgcg
 2281 tcggtatgcc gtggtccaga agcccgccca tgggacgtc cgcaaggcct cactcctggc
 2341 cgtctctacc ttgacagacc tccagccgta catgcgacag ttcgtggctc acctgcagga
 2401 gaccagcccc ctgagggatg ccgtcgtcat cagagcagag tcctccctga atgaggccag
 2461 cagtggcccg ttcgacgtct cctacgctt catgtgccac cagccgtgc gcactcagg
 2521 caagtccctac gtccagtgc gtccagtcga atcctctcca cgtctgtctg cgtgtgtctg
 2581 cagcctgtgc tacggcgaca tggagaacaa gctgtttgag gggattcggc gggacgggct
 2641 gctcctgcgt ttggtggatg atttcttgtt ggtgacacct cactcacc cgcgaaaaac
 2701 cttcctcagg accctggtcc gaggtgtccc ttagtatggc acttgcgga agctgcgga
 2761 gacagtgggt aacttccctg tagaagacga ggcctgggt ggacggctt ttgttcagat
 2821 gccggcccac ggcctattcc cctggtgcgg cctgctgctg gatacccga cctggagggt
 2881 gcagagcgac tactccagct atgcccggac ctccatcaga gccagtctca ccttcaaccg
 2941 cggcttcaag gctgggagga acatgcgtcg caaactctt ggggtcttgc ggctgaagt
 3001 tcacagcctg tttctggatt tgcaggtgaa cagcctccag acggtgtgca ccaacatcta
 3061 caagatcctc ctgctgcagg cgtacaggtt tcacgcagat cctgcgcgtc tcccatttca
 3121 tcagcaagtt tggaagaacc ccacattttt cctgcgcgtc agccctccct cggcctccct
 3181 ctgctactcc atcctgaaag ccaagaacgc agggatgtcg cctgtgccac agggcgccgc
 3241 cggccctctg ccctccgagg ccgtgcagtg gctgtgccac caagcattcc tgctcaagct
 3301 gactcgacac cgtgtcacct acgtgccact cctggggtca ctcaggacag cccagacgca
 3361 gctgagtcgg aagctcccgg ggacgacgct gactgccctg gaggccgag ccaaccggc
 3421 actgccctca gacttcaaga ccattcctgga ctgatggcca cccgcccaca gccaggccga
 3481 gacgacacac cagcagccct gtcacgcggg gctctacgtc ccagggaggg agggggcgcc
 3541 cacaccagg cccgaccgc tgggagctg aggcctgagt caggtgttgg cagaggcctg
 3601 catgtccggc tgaaggctga gtgtccggct gaggcctgag gctgttggc gccaagggct
 3661 gagtgtccag cacacctgcc gtcttcaact cccacaggc tggcgctcgg ctccacccca
 3721 gggccagctt ttctcacca ggagccggcg ttcactccc cactataggaa tagtccatcc
 3781 ccagattcgc cattgttccac gccctccctt gccttccacc cccaccatcc
 3841 aggtggagac cctgagaagg accctgggag cctctgggaat caaaggtgtg
 3901 ccctgtacac aggcgaggac cctgcacct gatgggggtc aaattggggg
 3961 gaggtgctgt gggagtaaaa tactgaatat atgagttttt cagttttgaa aaaaa

FIG. 16

MPRAPRCRAVRSLLRSHYREVLPLATFVRRRLGPQGWRLVQRGDP
 AAFRALVAQCLVCVPWDARPPPAAPSFRQVSCLKELVARVLQRL
 CERGAKNVLAFGFALLDARGGPPPEAFTTSVRSYLENTVTDALR
 GSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLY
 QLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPG
 ARRRGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRG
 PSDRGFCVVS PARPAEEATSLEGALSGTRHSHPSVGRQHHAGPP
 STSRPPRPWDTPCPPVYAETKHFYSSGDKEQLRPSFLLSSLRP
 SLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLEL
 LGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQSVAAPPEE
 EDTDPRRLVQLLRQHSSPWQVYGFVRACLRRLVPPGLWGSRHNE
 RRFLRNTKKFISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGC
 VPAAEHRLREEIILAKFLHWMMSVYVVELLRSFFYVTTETTFQKNR
 LFFYRKS VWSKLQSIGIRQHLKRVQLRELSAEVRQHREARPAL
 LTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKA
 LFSVLNRYERARRPGLLGASVLGLDDIHRWRTFVLRVRAQDPPP
 ELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQ
 KAAHGHVRKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDVVI
 EQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGI PQGS I
 LSTLLCSLCYGD MENKLFAGIRRDGLLLRLVDDFLLVTPHLTHA
 KTFRLTLVRGVPEYGCVVNLKRTVNVNFPVEDEALGGTAFVQMPA
 HGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR
 NMRRKLFGLRLKCHSLFLDLQVNSLQTVCTNIYKILLLOAYRF
 HACVLQLPFHQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL
 GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTVVPLLGSLRTAQ
 TQLSRKLPGTTLTALEAAANPALPSDFKTIID

FIG. 17

GGCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTT
 TTATGTACCGGAGACCACGTTTCAAAGAAGCAGGCTCTTTTCTACCGGAAGAGTGTCTG
 GAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGACGCTGCGGGAGCT
 GTCGGAAGCAGAGGTGAGGCAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACT
 CCGCTTCATCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGG
 AGCCAGAACGTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAAGGCAC
 GTTCAGCGTGCTCAACTACGAGCGGGCGCGGCCCGCCCTCCTGGGCGCCTCTGTGCT
 GGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTCGGGGCCCAGGA
 CCCGCCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC
 CCAGGACAGGCTCACGGAGGTGATCGCCAGCATCATCAAACCCAGAAACAGTACTGCGT
 CGCTCGGTATGCCGTGGTCCAGAAGGCCGCCATGGGCACGTCCGCAAGGCCCTTCAAGAG
 CCAGCTCCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCT
 GCAGCCTGTGCTACGGCGACATGGAGAACAAAGCTGTTTGCGGGGATTGCGCGGGACGGG
 TGCTCCTGCGTTTGGTGGATGATTTCTGTTGGTGACACCTCACCTCACCCACGCGAAAA
 CCTTCTCAGGACCTGGTCCGAGGTGTCCTGAGTATGGTGCGTGCTGGTGAACCTGCGGA
 AGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTGTTGAGA
 TGCCGGCCCCACGGCCTATTTCCCTGGTGCGGCTGCTGCTGGATACCCGGACCCCTGGAGG
 TGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACC
 GCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGT
 GTCACAGCCTGTTTCTGGATTTGCAAGGTGAACAGCCTCCAGACGGTGTCACCAACATCT
 ACAAGATCCTCCTGCTGCAGGCGTACAGGTTTACGCATGTGTGCTGCAGTCTCCATTTT
 ATCAGCAAGTTTGAAGAACCCACATTTTCTGCGCGTCATCTCTGACACGGCCTCCC
 TCTGCTACTCCATCCTGAAAGCCAAGAAGCAGGGATGTGCTGGGGGCCAAGGGCGCCG
 CCGGCC7TCTGCCCTCCGAGGCCGTGCACTGGCTGTGCCACCAAGCATTCCTGCTCAAGC
 TGACTCGACACCGTGTACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCAGACGC
 AGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGACGCCAACC
 CCGG CACTGCCCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCACAGCCAGGCCG
 AGAGCCAGTCCAGCAGCCCTGTGACGCCGGCTCTACGTCCCAGGGAGGGAGGGCGGC
 CCACACCCAGGCCGTGACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTGTTGGCCGAGGCCT
 GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGC
 TGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC
 AGGGCCAGCTTTTCTTCCACAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATC
 CCCAGATTGCGCATTTGTTACCCCTCGCCCTGCCCTCCTTTGCCCTTCCACCCCCACCATC
 CAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGT
 GCCCTGTACACAGGCGAGGACCTGACCTGGATGGGGGTCCCTGTGGGTCAAATTTGGGG
 GGAGGTGCTGTGGGAGTAAAAACTGAATATATGAGTTTTTCAGTTTTTG0AAAAA
 AAAAAAAAAAAAAAAAAA

FIG. 18

MetSerValTyrValValGluLeuLeuArgSerPhePhe
 TyrValThrGluThrThrPheGlnLysAsnArgLeuPhe
 PheTyrArgLysSerValTrpSerLysLeuGlnSerIle
 GlyIleArgGlnHisLeuLysArgValGlnLeuArgGlu
 LeuSerGluAlaGluValArgGlnHisArgGluAlaArg
 ProAlaLeuLeuThrSerArgLeuArgPheIleProLys
 ProAspGlyLeuArgProIleValAsnMetAspTyrVal
 ValGlyAlaArgThrPheArgArgGluLysArgAlaGlu
 ArgLeuThrSerArgValLysAlaLeuPheSerValLeu
 AsnTyrGluArgAlaArgArgProGlyLeuLeuGlyAla
 SerValLeuGlyLeuAspAspIleHisArgAlaTrpArg
 ThrPheValLeuArgValArgAlaGlnAspProProPro
 GluLeuTyrPheValLysValAspValThrGlyAlaTyr
 AspThrIleProGlnAspArgLeuThrGluValIleAla
 SerIleIleLysProGlnAsnThrTyrCysValArgArg
 TyrAlaValValGlnLysAlaAlaHisGlyHisValArg
 LysAlaPheLysSerHisValLeuArgProValProGly
 AspProAlaGlyLeuHisProLeuHisAlaAlaLeuGln
 ProValLeuArgArgHisGlyGluGlnAlaValCysGly
 AspSerAlaGlyArgAlaAlaProAlaPheGlyGly

FIG. 19

GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ¹met
 CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC ¹⁰
 pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
 CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC ¹⁰
 his tyr arg glu val leu pro leu ala thr phe val arg arg leu ³⁰
 CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG ³⁰
 gly pro gln gly trp arg leu val gln arg gly asp pro ala ala ⁴⁰
 GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT ⁴⁰
 phe arg ala leu val ala gln cys leu val cys val pro trp asp ⁶⁰
 TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC ⁶⁰
 ala arg pro pro pro ala ala pro ser phe arg gln val ser cys ⁷⁰
 GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC ⁷⁰
 leu lys glu leu val ala arg val leu gln arg leu cys glu arg ⁸⁰
 CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC ⁸⁰
 gly ala lys asn val leu ala phe gly phe ala leu leu asp gly ¹⁰⁰
 GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG ¹⁰⁰
 ala arg gly gly pro pro glu ala phe thr thr ser val arg ser ¹¹⁰
 GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC ¹¹⁰
 ala arg gly gly pro pro glu ala phe thr thr ser val arg ser ¹²⁰
 GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC ¹²⁰

FIG. 20A

tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
 TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

130
 trp gly leu leu leu arg arg val gly asp asp val leu val his
 TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

140
 leu leu ala arg cys ala leu phe val leu val ala pro ser cys
 CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

150
 ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
 GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

160
 thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
 ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

170
 leu gly cys glu arg ala trp asn his ser val arg glu ala gly
 CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

180
 val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
 GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

190
 ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
 AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

200
 ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
 GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

210
 ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
 GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

220
 val val ser pro ala arg pro ala glu glu ala thr ser leu glu
 GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

230
 gly ala leu ser gly thr arg his ser his pro ser val gly arg
 GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

240
 gln his his ala gly pro pro ser thr ser arg pro pro arg pro
 CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

250
 trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
 TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

260
 320
 330

FIG. 20B

340
 leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
 CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

350
 leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
 CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

360
 370
 glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
 GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

380
 arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
 CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

390
 400
 leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
 CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

410
 val leu leu lys thr his cys pro leu arg ala ala val thr pro
 GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

420
 430
 ala ala gly val cys ala arg glu lys pro gln gly ser val ala
 GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

440
 ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
 GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

450
 460
 leu arg gln his ser ser pro trp gln val tyr gly phe val arg
 CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

470
 ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
 GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

480
 490
 his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
 CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

500
 leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
 CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

510
 520
 met ser val arg asp cys ala trp leu arg arg ser pro gly val
 ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

530
 gly cys val pro ala ala glu his arg leu arg glu glu ile leu
 GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

540

FIG. 20C

550
 ala lys phe leu his trp leu met ser val tyr val val glu leu
 GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG
 560
 leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
 CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC
 580
 arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
 AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC
 590
 ile gly ile arg gln his leu lys arg val gln leu arg glu leu
 ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG
 610
 ser glu ala glu val arg gln his arg glu ala arg pro ala leu
 TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG
 620
 leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
 CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG
 640
 pro ile val asn met asp tyr val val gly ala arg thr phe arg
 CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC
 650
 arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
 AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG
 670
 phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
 TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG
 680
 gly ala ser val leu gly leu asp asp ile his arg ala trp arg
 GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC
 700
 thr phe val leu arg val arg ala gln asp pro pro pro glu leu
 ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG
 710
 tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
 TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC
 730
 gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
 CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG
 740
 asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
 AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC
 750

FIG. 20D

26/34

his gly his val arg lys ala phe lys ser his val leu arg pro
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA

760

val pro gly asp pro ala gly leu his pro leu his ala ala leu
GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG

770 780

gln pro val leu arg arg his gly glu gln ala val cys gly asp
CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT

790

ser ala gly arg ala ala pro ala phe gly gly OP
TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTTGGT

800 807

GACACCTCACCTCACCCACGCGAAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA
GTATGGCTGCGTGGTGAACCTTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGC
CCTGGGTGGCACGGCTTTTGTTCAGATGCCGGCCACGGCCTATTCCCCTGGTGCGGCCT
GCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTC
CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA
ACTCTTTGGGGTCTTGCGGCTGAAGTGTACAGCCTGTTTCTGGATTGTGAGGTGAACAG
CCTCCAGACGGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA
CGCATGTGTGCTGCAGCTCCCATTTTCATCAGCAAGTTTGGAAGAACCCACATTTTTCCT
GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG
GATGTGCTGCTGGGGGCCAAGGGCGCCGCCGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGCT
GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTACCTACGTGCCACTCCT
GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC
TGCCCTGGAGGCCGAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG
ATGGCCACCCGCCCCACAGCCAGGCCGAGAGCAGACACCAGCAGCCCTGTACGCCGGGCT
CTACGTCCCAGGGAGGGAGGGGCGGCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG
CCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG
GCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCC
CACAGGCTGGCGCTCGGCTCCACCCCAGGGCCAGCTTTTCTCACCAGGAGCCCGGCTTC
CACTCCCCACATAGGAATAGTCCATCCCCAGATTGCGCATTGTTACCCCCTCGCCCTGCC
CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC
TGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT
GGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATG
AGTTTTTTCAGTTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIG. 20E

+

FIG. 21A

+

28/34

4321 GGTGTTTTTAAGCCAATNANAAAATTTTTTNATGTTGTTNNNNNNNNNNNNNNNNNNNN
CCACAAAATTCGGTTANTNTTTTAAAAAANTACAACAAANNNNNNNNNNNNNNNNNNNNN

4381 NNN
NN

4441 NNN
NN

4501 NNN
NN

4561 NNN
NN

4621 NNN
NN

4681 NNN
NN

4741 NNN
NN

4801 NNN
NN

4861 NNN
NN

4921 NNN
NN

4981 NNN
NN

5041 NGCCANGRAGGGGGCCAGGTTCCAANTTCCCAACCKTTTWTGGARGGACNGCCCCCAGGG
NCGGTNCYTCCCCCGGTCCAAGGTTNAAGGGTTGGMAAAAWCCTYCCTGNCGGGGGTCCC

5101 GGGGATRAACAGANTNGGGGKGGTWGGGTTNAKGGTGGGAACNCCTTNGCGCCTGGAG
CCCCTAYTTGTCTNANCCCCMCCAWCCCAANTMCCACCCTTGNGGAANCGSCGGACCTC

5161 AACGTGCAAAGAGGAAATGAAGGGCCTGKGTCAAGGAGCCCAAGTNGGCGGGGRAGTTTG
TTGCACGTTTCTCCTTTACTTCCCGGACMCAGTTCTCGGGTTCANCCGCCCCYTCAAAC

5221 CAGGGAGGCACTCCGGGGAGGTCCSGCGTGCCCGTCCAAGGGAGCAATGCGTCCTTCGGG
GTCCCTCCGTGAGGCCCTCCAGGSCGCACGGGCAGGTTCCCTCGTTACGCAGGAAGCCC

5281 TTCGTCCCCAWGCCGCTCTACGCGCTYCCGTCTTCCCTTTCAGGTTCCGGCATTTCGTG
AAGCAGGGGTWCGGCGCAGATGCGCGGARGGCAGGAGGGGAAGTGCAAGGCCGTAAGCAC

5341 GTGCCCCGAGCCCGACGCCCCGCGTCCGGACCTGGAGGCAGCCCTGGGTCTCCGGATCAG
CACGGGCCTCGGGCTGCGGGGCGCAGGCCTGGACCTCCGTGCGGACCCAGAGGCCTAGTC

5401 GCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCAGGGCCTCCACATCATGGCCCCCT
CGGTCGCCGGTTTCCAGCGGCGTGCGTGGACAAGGGTCCCGGAGGTGTAGTACCGGGGA

FIG. 21B

+

+

29/34

5461 CCCTCGGGTTACCCACAGCCTAGGCCGATTTCGACCTCTCTCCGCTGGGGCCCTCGCCT
GGGAGCCCAATGGGGTGTTCGGATCCGCGCTAAGCTGGAGAGAGGCGACCCCGGGAGCGGA

Sp1

5521 GCGTCCCTGCACCCTGGGAGCGCGAGCGGCGCGGGCGGGGAAGCGCGGCCCATACCC
CCGACGGGACGTGGGACCCTCGCGCTCGCCGCGCGCCCGCCCTTCGCGCCGGGTATGGG

5581 CCGGGTCCGCCCAGGAGCAGCTGCGCTGTGCGGGCCAGGCCGGGCTCCAGTGGATTTCG
GGCCAGGCGGGCCTTCGTGACGCGACAGCCCCGGTCCGGCCCCGAGGGTCACCTAAGCG

Topo_II_cleavage_site

5641 GGGCACAGACGCCCAGGACCGCGCTTCCACGTGGCGGAAGGACTGGGGACCCGGGCACC
CCCGTGTCTGCGGGTCCTGGCGCGAAGGGTGCACCGCCTTCCTGACCCCTGGGCCCGTGG

E2F

5701 CGTCCTGCCCCCTTCACCTTCCAGCTCCGCTTCTTCCGCGCGGACCCGGCCCCGTCCCGAA
GCAGGACGGGGAAGTGGAAGGTGAGGCGAAGAAGGCGCGCCTGGGCCGGGGCAGGGCTT

E

5761 CCCTTCCCAGGTCCCGGGCCAGCCCCCTTCCGGGCCCTCCCAGCCCCCTCCCCTTCTTTTC
GGGAAGGGTCCAGGGCCGGGTGCGGGAAGGCCCGGGAGGGTTCGGGGAGGGGAAGGAAAAG

Sp1

=====

2F

NFkB

h

5821 CGCGGCCCCCGCCCTCTCCTTCGCGGCGCGAGTTTCAGGCAGCGCTGCGTCTGCTGCGCA
GCGCCGGGGCGGGAGAGGAAGCGCCGCGCTCAAAGTCCGTTCGCGACGCAGGACGACGCGT

5860

ECO47III

5875

FSP1

TRT5'

*****>

5881 CGTGGAAGCCCTGGCCCCGGCCACCCCGCGATGCCGCGCGCTCCCCGCTGCCGAGCCG
GCACCCTTCGGGACCGGGGCCGGTGGGGGCGCTACGGCGCGCGAGGGGCGACGGCTCGGC

5941 TGCGTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTCTGTGCGGC
ACGCGAGGGACGACGCTCGGTGATGGCGCTCCACGACGGCGACCGGTGCAAGCACGCCG

5953

FSP1

6001 GCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGACCCGGCGGCTTTCCGCGCGC
CGACCCCGGGGTCCCGACCGCCGACCACGTGCGCCCCCTGGGCCGCGGAAAGGCGCGCG

6061 TGGTGGCCCAGTGCCCTGGTGTGCGTGCCCTGGGACGCACGGCCGCCCCCGCCGCCCT
ACCACGGGTACGGACCACACGCACGGGACCCTGCGTGCCGGCGGGGGCGGCGGGGGA

NFkB

=====

FIG. 21C

+

+

30/34

```
*****
6121 CCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGG
    GGAAGGCGGTCCACCCGAGGGGCCCCAGCCGACGGCCGACCCAACTCCCGCCGGCCCC

                                         Topo_II_cleavage_s
                                         ::::::::::::::
                                         NFkB
                                         =====

Intron1
*****>
6181 GGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTCAGGGCGCTTCCCCCGCAGGTGTC
    CCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAGTCCCGCGAAGGGGGCGTCCACAG

ite
:

6241 CTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCGGAAGAA
    GACGGACTTCTCTGACCACCGGGCTCACGACGTCTCCGACACGCTCGCGCCGCGCTTCTT

6301 CGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCGCGGGGGCCCCCGAGGCCTT
    GCACGACCGGAAGCCGAAGCGCGACGACCTGCCCCGGGCGCCCCGGGGGGGCTCCGGAA

6361 CACCACCAGCGTGCGCAGCTACCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGG
    GTGGTGGTCGCACGCGTCGATGGACGGGTTGTGCCACTGGCTGCGTGACGCCCCCTCGCC
    ^
    6372
    FSP1

6421 GGCCTGGGGGCTGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTACCTGCTGGCACG
    CCGACCCCCGACGACGACGCGGCGCACCCGCTGCTGCACGACCAAGTGGACGACCGTGC

6481 CTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCCGCT
    GACGCGCGAGAAACACGACCACCGAGGGTCGACGCGGATGGTCCACACGCCCGGCGGCGA

6541 GTACCAGCTCGGCGCTGCCACTCAGGCCCCGGCCCCCGCCACACGCTAGTGGACCCGAAG
    CATGGTCGAGCCGCGACGGTGAGTCCGGGCGGGGGCGGTGTGCGATCACCTGGGGCTTC

6601 GCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCGGGGTCCCCCTGGG
    CGCAGACCCTACGCTTGCCCCGACCTTGGTATCGCAGTCCCTCCGCCCCAGGGGGACCC

6661 CCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCC
    GGACGGTCGGGGCCCACGCTCCTCCGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGG

6721 CAAGAGGCCCAGGCGTGGCGCTGCCCCTGAGCCGAGCGGACGCCCCTTGGGCAGGGGTC
    GTTCTCCGGGTCCGCACCGCGACGGGACTCGGCCTCGCCTGCGGGCAACCCGTCCCCAG

6781 CTGGGCCCCACCCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTTCTGTGTGGTGTACCC
    GACCCGGGTGGGCCCCGTCTGCGCACCTGGCTCACTGGCACCAAAGACACACCACAGTGG

6841 TGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTC
    ACGGTCTGGGCGGCTTCTTCGGTGGAGAAACCTCCACGCGAGAGACCGTGCGCGGTGAG

6901 CCACCCATCCGTGGGCCGCCAGCACCACGCGGGCCCCCATCCACATCGCGGCCACCACG
    GGTGGGTAGGCACCCGGCGGTCTGTGGTGCGCCCGGGGGTAGGTGTAGCGCCGGTGGTGC
```

FIG. 21D

+

+

31/34

6961 TCCCTGGGACACGCCTTGTCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTC
AGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAG
7021 AGGCGACAAGGAGCAGCTGCGGCCCTCCTTCTACTCAGCTCTCTGAGGCCCAGCCTGAC
TCCGCTGTTCTCGTCGACGCCGGGAGGAAGGATGAGTCGAGAGACTCCGGGTCGGACTG
7081 TGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCAGGCCCTGGATGCCAGGGAC
ACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTG
7141 TCCCCGAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGGCCCTGTTTCTGGA
AGGGGCGTCCAACGGGGCGGACGGGGTCGCGATGACCGTTTACGCCGGGGACAAAGACCT

7167
ECO47III

7201 GCTGCTTGGGAACCACGCGCAGTGCCCCCTACGGGGTGCTCCTCAAGACGCACTGCCCGCT
CGACGAACCCCTGGTGCGCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGGCGA
7261 GCGAGCTGCGGTACCCCAGCAGCCGGTGTCTGTGCCCCGGGAGAAGCCCCAGGGCTCTGT
CGCTCGACGCCAGTGGGGTCGTCGGCCACAGACACGGGCCCTCTTCGGGGTCCCAGACA
7321 GGCGGCCCCCGAGGAGGAGGACACAGACCCCCGTGCTTGGTGCAGCTGCTCCGCCAGCA
CCGCCGGGGGCTCCTCCTCTGTGTCTGGGGGACGCGGACCACGTGACGAGGCGGTCTGT
7381 CAGCAGCCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCTGCGCCGGCTGGTGCCCC
GTCGTGCGGGACCGTCCACATGCCGAAGCACGCCCGGACGGACGCGGCCGACCACGGGGG
7441 AGGCCTCTGGGGCTCCAGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCAT
TCCGGAGACCCCGAGGTCCGTGTTGCTTGCGGCGAAGGAGTCCTTGTTGTTCTTCAAGTA
7501 CTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCG
GAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCTCTGACTGCACCTTCTACTCGCACGC

7561 GGA CTGCGCTTGGCTGCGCAGGAGCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCAGG
CCTGACGCGAACCAGCGCTCCTCGGGTCCACTCCTCCACCACCGGCAGCTCCCGGGTCC

7575
FSP1

Intron2

7621 CCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCCCTGGTCTCTCT
GGGGTCTCGACTTACGTATCCCCGAGTCTTTTCCCCGTCCGTCTCGGGACCAGGAGGA

7681 GTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTGAGTGGACACGGTG
CAGAGGTAGCAGTGCACCCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGCCAC

**>
7741 ATCGAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTTCGTAATCATGGTCATA
TAGCTCCAGCTGAGATCTCCTAGGGGGCCATGGCTCGAGCTTAAGCATTAGTACCAGTAT

7747
SAL1

FIG. 21E

+

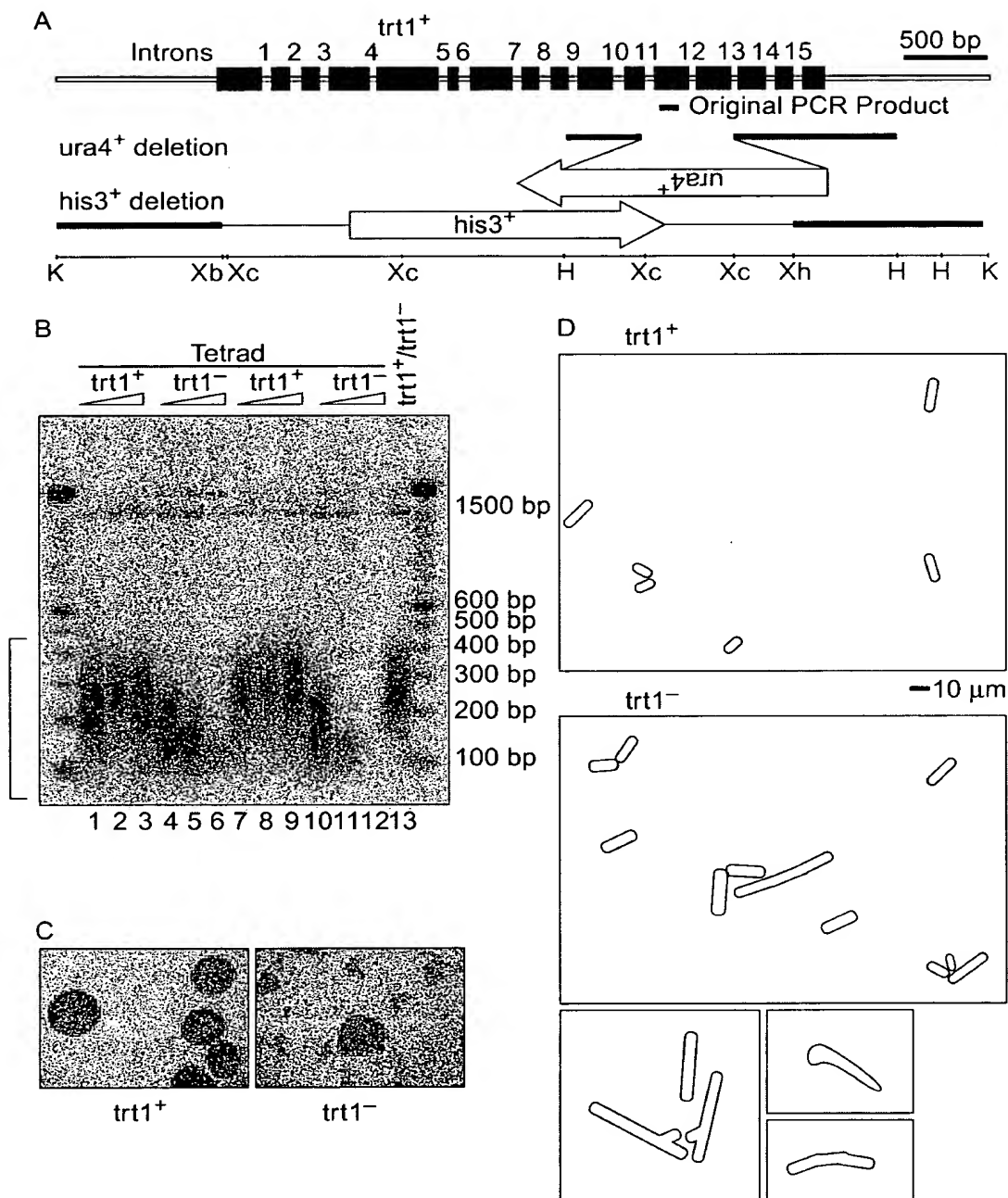


FIG. 22

gccaagttcctgcaactggctgatgagtgtgtacgtcgtcgagctgctcaggtctttcttt
 tatgtcacggagaccacgtttcaaaagaacaggctctttttctaccggaagagtgtctgg
 agcaagttgcaaagcattggaatcagacagcacttgaagaggggtgcagctgcgggacgtg
 tcggaagcagaggtcaggcagcatcggaagccaggcccgccctgctgacgtccagactc
 cgcttcatccccaagcctgacgggctgcggccgattgtgaacatggactacgtcgtggga
 gccagaacgttccgcagagaaaagagggccgagcgtctcacctcgaggggtgaaggcactg
 ttcagcgtgctcaactacgagcgggcgcg

FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG
 ACCAGCCCGCTGAGGGATGCCGTCGTTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC
 AGCAGTGGCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACACGCCGTGCGCATC
 AGGGGCAAGTC

FIG. 24

+

34/34

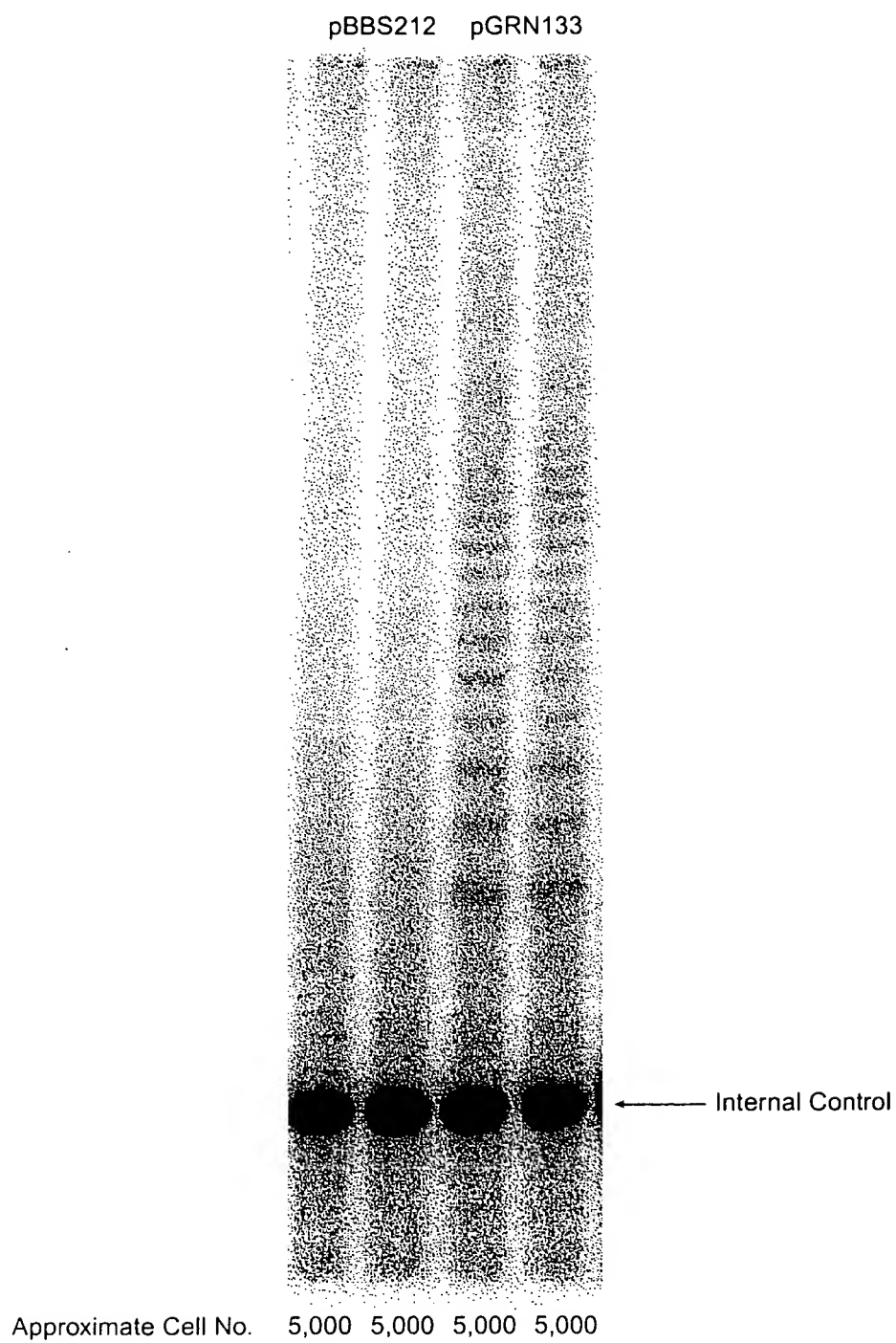


FIG. 25

+

+

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA 1529
 129 D L V S T F P N Y L I S I L E S K N W Q 148
 1530 CTT TTG TTA GAA AT gtaataaccgggtaagatgttgccgacacttggaacaagactgacaagtatag T ATC GGC 1601
 149 L L L E I I G 155
 1602 AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661
 156 S D A M H Y L L S K G S I F E A L P N D 175
 1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721
 176 N Y L Q I S G I P L F K N N V F E T V 195
 1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781
 196 S K K R K R T I E T S I T Q N K S A R K 215
 1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841
 216 E V S W N S I S I S R F S I F Y R S S Y 235
 1842 AAG AAG TTT AAG CAA G gtaactaactgttattccttcataactaatttttag AT CTA TAT TTT AAC 1907
 236 K K F K Q D L Y F N 245
 1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967
 246 L H S I C D R N T V H M W L Q W I F P R 265
 1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027
 266 Q F G L I N A F Q V K Q L H K V I P L V 285
 2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087
 286 S Q S T V V P K R L L K V Y P L I E Q T 305
 2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147
 306 A K R L H R I S L S K V Y N H Y C P Y I 325
 2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207
 326 D T H D D E K I L S Y S L K P N Q V F A 345
 2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267
 346 F L R S I L V R V F P K L I W G N Q R I 365

FIG. 15B



+

+

2268 TTT GAG ATA ATA TTA AAA G gtaattgtataaaattattaccactaacaagattttaccag AC CTC GAA ACT 2336
 366 F E I I L K D L E T 375
 2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
 376 F L K L S R Y E S F S L H Y L M S N I K 395
 2397 gtaatatgccaaattttttaccatttaataacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465
 396 I S E I E W L V L G 405
 2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
 406 K R S N A K M C L S D F E K R K Q I F A 425
 2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
 426 E F I Y W L Y N S F I I P I L Q S F F Y 445
 2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
 446 I T E S S D L R N R T V Y F R K D I W K 465
 2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
 466 L L C R P F I T S M K M E A F E K I N E 485
 2706 gtaatttaaagtattttttgcaaaaagtaataattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
 486 N N V R M D T Q K T 495
 2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835
 496 T L P P A V I R L L P K K N T F R L I T 515
 2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtatttaatttttggtcatcaatgtactttacttctaattctatta 2906
 516 N L R K R F L I K 524
 2907 ttagcag ATG GGT TCA AAC AAA AAT ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
 525 M G S N K K M L V S T N Q T L R P V 542
 2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
 543 A S I L K H L I N E S S G I P F N L E 562
 3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaat 3088
 563 V Y M K L L T F K K D L L K H R M F G 581

FIG. 15C

+

+

```

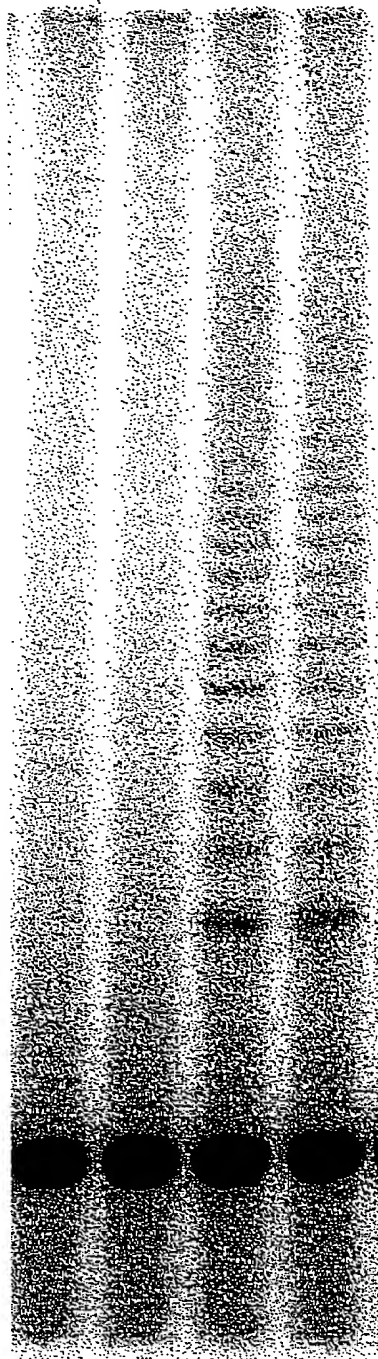
3089 tatataatgcgcgattcctcattatttaattttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 3155
582      R K K Y F V R I D I 591
3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 3215
592 K S C Y D R I K Q D L M F R I V K K L 611
3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3275
612 K D P E F V I R K Y A T I H A T S D R A 631
3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttatttttccattggaattttttaacaa 3343
632 T K N F V S E A F S Y F 643
3344 attcttttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3405
644      D M V P F E K V V Q L L S M K T 659
3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3465
660 S D T L F V D F V D Y W T K S S E I F 679
3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgtgaattgtaataaca 3532
680 K M L K E H L S G H I V K
3533 ctaatgaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA 3593
693      I G N S Q Y L Q K V G I P Q G S 708
3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3653
709 I L S S F L C H F Y M E D L I D E Y L S 728
3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3713
729 F T K K K G S V L L R V V D D F L F I T 748
3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgctgtcattcc 3777
749 V N K K D A K K F L N L S L R G 764
3778 taagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3840
765      F E K H N F S T S L E K T V 778
3841 ATA AAC TTT GAA AAT AGT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA 3900
779 I N F E N S N G I I N N T F F N E S K K 798

```

FIG. 15D

+

pBBS212 pGRN133



← Internal Control

Approximate Cell No. 5,000 5,000 5,000 5,000

FIG. 25

08/9/295/

gccaagttcctgcactggctgatgagtgtgtacgtcgtcagctgctcaggtctttcttt
tatgtcacggagaccagtttcaaaagaacaggctctttttctaccggaagagtgtctgg
agcaagttgcaaagcattggaatcagacagcacttgaagaggggtgcagctgcgggacgtg
tcggaagcagaggtcaggcagcatcgggaagccaggccccgcctgctgacgtccagactc
cgcttcatccccaagcctgacgggctgcggccgattgtgaacatggactacgtcgtggga
gccagaacgttccgcagagaaaaagagggccgagcgtctcacctcgaggggtgaaggcactg
ttcagcgtgctcaactacgagcgggcgcg

FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG
ACCAGCCCCGTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC
AGCAGTGGCCTCTTCGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATC
AGGGGCAAGTC

FIG. 24

08/9/2951

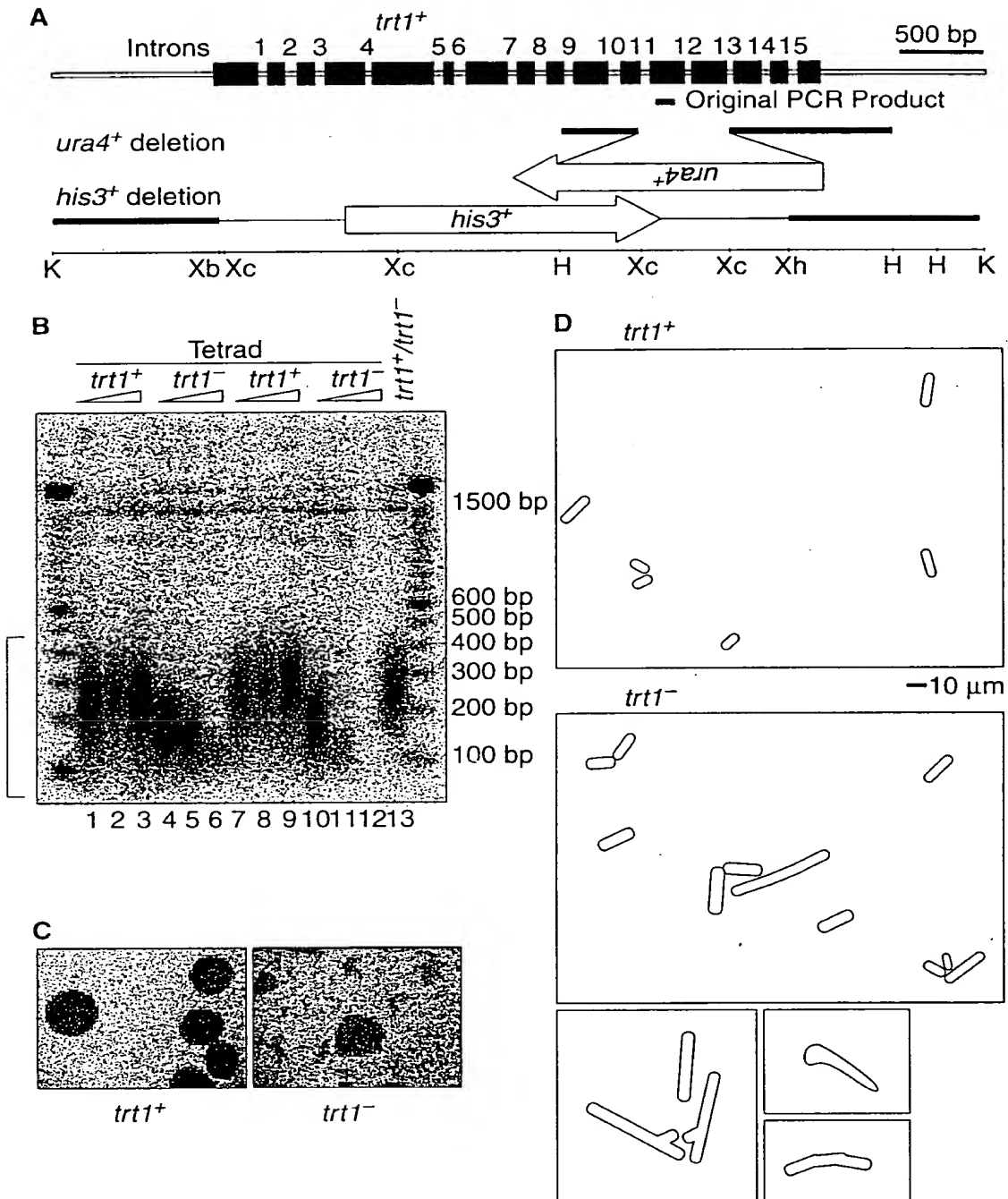


FIG. 22

6961 TCCCTGGGACACGCCTTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTC
AGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTCGTGAAGGAGATGAGGAG

7021 AGGCGACAAGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCAGCCTGAC
TCCGCTGTTCTCGTCGACGCCGGGAGGAAGGATGAGTCGAGAGACTCCGGGTGCGACTG

7081 TGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCAGGCCCTGGATGCCAGGGAC
ACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCAAGGTCCGGGACCTACGGTCCCTG

7141 TCCCCGCAGGTTGCCCCGCTGCCCCAGCGCTACTGGCAAATGCGGGCCCTGTTTCTGGA
AGGGGCGTCCAACGGGGCGGACGGGGTTCGCGATGACCGTTTACGCCGGGGACAAAGACCT

7167
ECO47III

7201 GCTGCTTGGAACCACGCGCAGTGCCCCCTACGGGGTGCTCCTCAAGACGCACTGCCCCGCT
CGACGAACCCTTGGTGC GCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGGCGA

7261 GCGAGCTGCGGTACCCCCAGCAGCCGGTGTCTGTGCCCCGGGAGAAGCCCCAGGGCTCTGT
CGCTCGACGCCAGTGGGGTTCGTGCGCCACAGACACGGGGCCCTCTTCGGGGTCCCGAGACA

7321 GGCGGCCCCCGAGGAGGAGGACACAGACCCCCGTGCGCTGGTGCAGCTGCTCCGCCAGCA
CCGCCGGGGGCTCCTCCTCCTGTGTCTGGGGGCAGCGGACCACGTGACGAGGCGGTTCGT

7381 CAGCAGCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCTGCGCCGGCTGGTGCCCCC
GTCGTGCGGGACCGTCCACATGCCGAAGCACGCCCGGACGGACGCGGCCGACCACGGGGG

7441 AGGCCTCTGGGGCTCCAGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCAT
TCCGGAGACCCCGAGGTCCGTGTTGCTTGCGGCGAAGGAGTCCTTGTGGTTCTTCAAGTA

7501 CTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCG
GAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCTCGACTGCACCTTCTACTCGCACGC

7561 GGA CTGCGCTTGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCCAGG
CCTGACGCGAACCAGCGCTCCTCGGGTCCACTCCTCCACCACCGGCAGCTCCCGGGTCC

7575
FSP1

Intron2

7621 CCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGCAGGCAGAGCCCTGGTCTCCT
GGGGTCTCGACTTACGTCATCCCCGAGTCTTTTCCCCCGTCCGTCTCGGGACCAGGAGGA

7681 GTCTCCATCGTCACGTGGGCACACGTGGCTTTTTCGCTCAGGACGTCGAGTGGACACGGTG
CAGAGGTAGCAGTGCACCCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGCCAC

**>

7741 ATCGAGGTGCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTTCGTAATCATGGTCATA
TAGCTCCAGCTGAGATCTCCTAGGGGCCCATGGCTCGAGCTTAAGCATTAGTACCAGTAT

7747
SAL1

08/9/2951

 6121 CCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGG
 GGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGAGGCCGACCCCAACTCCCGCCGGCCCC

Topo_II_cleavage_s

::::::::::::::::::

NFkB

=====

Intron1

*****>

6181 GGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTCAGGGCGCTTCCCCCGCAGGTGTC
 CCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAGTCCCGCGAAGGGGGCGTCCACAG

ite

:

6241 CTGCCTGAAGGAGCTGGTGGCCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCCGGAAGAA
 GACGGACTTCTCTGACCACCGGGCTCACGACGTCTCCGACACGCTCGCGCCGCGCTTCTT

6301 CGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCCGCGGGGGCCCCCGAGGCCTT
 GCACGACCGGAAGCCGAAGCGCGACGACCTGCCCGGGCGCCCCGGGGGGGCTCCGGAA

6361 CACCACCAGCGTGCGCAGCTACCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGG
 GTGGTGGTCGCACGCGTCGATGGACGGGTGTGCCACTGGCTGCGTGACGCCCCCTCGCC

6372

FSP1

6421 GGCGTGGGGGCTGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTACCTGCTGGCAGC
 CCGCACCCCCGACGACGACGCGGCGCACCCGCTGCTGCACGACCAAGTGGACGACCGTGC

6481 CTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCGCT
 GACGCGCGAGAAACACGACCACCGAGGGTCGACGCGGATGGTCCACACGCCCGGCGGCGA

6541 GTACCAGCTCGGCGCTGCCACTCAGGCCCCGCCCCGCCACACGCTAGTGGACCCCGAAG
 CATGGTCGAGCCGCGACGGTGAGTCCGGGCGGGGGCGGTGTGCGATCACCTGGGGCTTC

6601 GCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCTGGG
 CGCAGACCCTACGCTTGCCCGGACCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGACCC

6661 CCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGCGAGTGCCAGCCGAAGTCTGCCGTTGCC
 GGACGGTCGGGGCCACGCTCCTCCGCGCCCCCGTCACGGTCGGCTTCAGACGGCAACGG

6721 CAAGAGGCCAGGCGTGCGCTGCCCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGGGTC
 GTTCTCCGGGTCCGCACCGCGACGGGACTCGGCCTCGCCTGCGGGCAACCCGTCCCCAG

6781 CTGGGCCCCACCGGGCAGGACGCGTGACCGAGTGACCGTGGTTTCTGTGTGGTGTACAC
 GACCGGGTGGGGCCGTCCTGCGCACCTGGCTCACTGGCACCAAGACACACCACAGTGG

6841 TGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTC
 ACGGTCTGGGCGGCTTCTTCGGTGGAGAAACCTCCCACGCGAGAGACCGTGCGCGGTGAG

6901 CCACCCATCCGTGGGCCGCCAGCACACGCGGGCCCCCATCCACATCGCGGCCACCACG
 GGTGGGTAGGCACCCGGCGGTGCTGGTGGCGCCGGGGGTAGGTGTAGCGCCGGTGGTGC

FIG. 21
 (CONTINUED)

08/9/2951

5461 CCCTCGGGTTACCCACAGCCTAGGCCGGATTTCGACCTCTCTCCGCTGGGGCCCTCGCCT
GGGAGCCCAATGGGGTGTTCGGATCCGGCCTAAGCTGGAGAGAGGCGACCCGGGAGCGGA

Sp1

5521 GGCGTCCCTGCACCCTGGGAGCGCGAGCGGCGCGGGCGGGGAAGCGCGGCCATACCC
CCGACGGGACGTGGGACCCTCGCGCTCGCCGCGCGCCCGCCCTTCGCGCCGGGTATGGG

5581 CCGGGTCCGCCCCGAAGCAGCTGCGCTGTTCGGGGCCAGGCCGGGCTCCAGTGGATTTCG
GGCCCAGGCGGGCCTTCGTTCGACGCGACAGCCCCGGTCCGGCCCCAGGGTCACCTAAGCG

Topo_II_cleavage_site

5641 GGGCACAGACGCCCAGGACCGCGCTTCCACGTGGCGGAAGGACTGGGGACCCGGGCACC
CCCGTGTCTGCGGGTCCTGGCGCGAAGGGTGCACCGCCTTCCTGACCCCTGGGCCCCGTGG

E2F

5701 CGTCCTGCCCCCTTACCTTCCAGCTCCGCTTCTTCCGCGCGGACCCGGCCCCGTCCCGAA
GCAGGACGGGGAAGTGGAAGGTCGAGGCGAAGAAGGCGCGCCTGGGCCGGGGCAGGGCTT

E

5761 CCCTTCCCAGGTCCCGGCCAGCCCCCTTCCGGGCCCTCCAGCCCCCTCCCCTTCTTTTC
GGGAAGGGTCCAGGGCCGGGTCGGGGAAGGCCCGGGAGGGTCGGGGAGGGGAAGGAAAAG

Sp1

=====

2F

NFkB

h

5821 CGCGGCCCCGCCCTCTCCTTCGCGGCGCGAGTTTCAGGCAGCGCTGCGTCCTGCTGCGCA
GCGCCGGGGCGGGAGAGGAAGCGCCGCGCTCAAAGTCCGTCGCGACGCAGGACGACGCGT

5860

ECO47III

5875

FSP1

TRT5'

*****>

5881 CGTGGGAAGCCCTGGCCCCGGCCACCCCGCGATGCCGCGCGCTCCCCGCTGCCGAGCCG
GCACCCTTCGGGACCGGGCCGGTGGGGGCGCTACGGCGCGCGAGGGGCGACGGCTCGGC

5941 TCGCTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTTCGTGCGGC
ACGCGAGGGACGACGCGTCGGTGATGGCGCTCCACGACGGCGACCGGTGCAAGCACGCCG

5953

FSP1

6001 GCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACCCGGCGGCTTTCGCGCGC
CGGACCCCGGGGTCCCGACCGCCGACCACGTTCGCGCCCCCTGGGCCGCCGAAAGGCGCGC

6061 TGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGCACGGCCGCCCCCGCGCCCCCT
ACCACGGGTACGGACCACACGCACGGGACCCTGCGTGCCGGCGGGGGGCGCGGGGGA

NFkB

=====

08/9/2951

4321 GGTGTTTTTAAGCCAATNANAAAAATTTTTTNATGTTGTTNNNNNNNNNNNNNNNNNNNNNN
 CCACAAAAATTCGGTTANTNTTTTAAAAAANTACAACAAANNNNNNNNNNNNNNNNNNNNNN
 4381 NNN
 NNN
 4441 NNN
 NNN
 4501 NNN
 NNN
 4561 NNN
 NNN
 4621 NNN
 NNN
 4681 NNN
 NNN
 4741 NNN
 NNN
 4801 NNN
 NNN
 4861 NNN
 NNN
 4921 NNN
 NNN
 4981 NNN
 NNN
 5041 NGCCANGRAGGGGGCCAGGTTCCAANTTCCCAACCKTTTWWGGARGGACNGCCCCAGGG
 NCGGTNCYTCCCCCGGTCCAAGGTTNAAGGGTTGMAAAAWCCTYCCTGNCGGGGGTCCC
 5101 GGGGATRAACAGANTNGGGGGKGGTWWGGTTNAKGGTGGGAACNCCTTNGCGCCTGGAG
 CCCCAYTTGTCTNANCCCCMCCAWCCCAANTMCCACCCTTGNGGAANCGSCGGACCTC
 5161 AACGTGCAAAGAGGAAATGAAGGGCCTGKGTCAAGGAGCCCAAGTNGGCGGGGRAGTTTG
 TTGCACGTTTCTCCTTTTACTTCCCGGACMCAGTTCTCGGGTTCANCCGCCCCYTCAAAC
 5221 CAGGGAGGCACTCCGGGGAGGTCCSGCGTGCCCGTCCAAGGGAGCAATGCGTCCCTTCGGG
 GTCCCTCCGTGAGGCCCTCCAGGSCGCACGGGCAGGTTCCTCGTTACGCAGGAAGCCC
 5281 TTCGTCCCCAWGCCGCGTCTACGCGCCTYCCGTCCCTCCCTTCACGTTCGGGCATTCGTG
 AAGCAGGGGTWCGGCGCAGATGCGCGGARGGCAGGAGGGGAAGTGCAAGGCCGTAAGCAC
 5341 GTGCCCCGAGCCCCGACGCCCCGCGTCCGGACCTGGAGGCAGCCCTGGGTCTCCGGATCAG
 CACGGGCCTCGGGCTGCGGGGCGCAGGCCTGGACCTCCGTGCGGACCCAGAGGCCTAGTC
 5401 GCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCAGGGCCTCCACATCATGGCCCCCT
 CGGTGCCCGGTTTCCCAGCGGCGTGCGTGGACAAGGGTCCCGGAGGTGTAGTACCGGGGA

FIG. 21
(CONTINUED)

08/9/2951

08/9/295/

08/9/295/

his gly his val arg lys ala phe lys ser his val leu arg pro
 CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA
 760
 val pro gly asp pro ala gly leu his pro leu his ala ala leu
 GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG
 770
 gln pro val leu arg arg his gly glu gln ala val cys gly asp
 CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT
 790
 ser ala gly arg ala ala pro ala phe gly gly OP
 TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTTGGT
 800 807
 GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA
 GTATGGCTGCGTGGTGAACCTTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGC
 CCTGGGTGGCAGGCTTTTGTTCAGATGCCGGCCACGGCCTATTCCCCTGGTGCGGCCT
 GCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCCGACCTC
 CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA
 ACTCTTTGGGGTCTTGCGGCTGAAGTGTACAGCCTGTTTCTGGATTTGCAGGTGAACAG
 CCTCCAGACGGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA
 CGCATGTGTGCTGCAGCTCCCATTTTCATCAGCAAGTTTGGAAGAACCCACATTTTTCCT
 GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG
 GATGTGCTGTTGGGGCCAAAGGGCGCCGCCGCTCTGCCCTCCGAGGCCGTGCAGTGGCT
 GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTACCTACGTGCCACTCCT
 GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC
 TGCCCTGGAGGCCGAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG
 ATGGCCACCCGCCCACAGCCAGGCCGAGAGCAGACACCAGCAGCCCTGTCACGCCGGGCT
 CTACGTCCCAGGGAGGGAGGGGCGGCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG
 CCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG
 GCCTGAGCGAGTGTCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCC
 CACAGGCTGGCGCTCGGCTCCACCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTC
 CACTCCCCACATAGGAATAGTCCATCCCCAGATTGCGCCATTGTTTACCCCTCGCCCTGCC
 CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC
 TGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT
 GGGGGTCCCTGTGGGTCAAATTGGGGGGAGGTGCTGTGGGAGTAAATACTGAATATATG
 AGTTTTTTCAGTTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIG. 20
 (CONTINUED)

08/912951

25/34

550
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

560
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

580
arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC

590
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

610
ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

620
leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

640
pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

650
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

670
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG

680
gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

700
thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

710
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

730
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

740
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

750

FIG. 20
(CONTINUED)

08/9/2951

340
 leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
 CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

350
 leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
 CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

370
 glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
 GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

380
 arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
 CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

400
 leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
 CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

410
 val leu leu lys thr his cys pro leu arg ala ala val thr pro
 GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

430
 ala ala gly val cys ala arg glu lys pro gln gly ser val ala
 GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

440
 ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
 GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

460
 leu arg gln his ser ser pro trp gln val tyr gly phe val arg
 CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

470
 ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
 GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

490
 his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
 CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

500
 leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
 CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

520
 met ser val arg asp cys ala trp leu arg arg ser pro gly val
 ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

530
 gly cys val pro ala ala glu his arg leu arg glu glu ile leu
 GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

540

FIG. 20
(CONTINUED)

130
 tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
 TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

140
 trp gly leu leu leu arg arg val gly asp asp val leu val his
 TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

160
 leu leu ala arg cys ala leu phe val leu val ala pro ser cys
 CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

170
 ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
 GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

190
 thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
 ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

200
 leu gly cys glu arg ala trp asn his ser val arg glu ala gly
 CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

220
 val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
 GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

230
 ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
 AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

250
 ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
 GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

260
 ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
 GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

280
 val val ser pro ala arg pro ala glu glu ala thr ser leu glu
 GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

290
 gly ala leu ser gly thr arg his ser his pro ser val gly arg
 GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

310
 gln his his ala gly pro pro ser thr ser arg pro pro arg pro
 CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

320
 trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
 TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

330

FIG. 20
(CONTINUED)

08/9/2951

22/34

MetSerValTyrValValGluLeuLeuArgSerPhePhe
TyrValThrGluThrThrPheGlnLysAsnArgLeuPhe
PheTyrArgLysSerValTrpSerLysLeuGlnSerIle
GlyIleArgGlnHisLeuLysArgValGlnLeuArgGlu
LeuSerGluAlaGluValArgGlnHisArgGluAlaArg
ProAlaLeuLeuThrSerArgLeuArgPheIleProLys
ProAspGlyLeuArgProIleValAsnMetAspTyrVal
ValGlyAlaArgThrPheArgArgGluLysArgAlaGlu
ArgLeuThrSerArgValLysAlaLeuPheSerValLeu
AsnTyrGluArgAlaArgArgProGlyLeuLeuGlyAla
SerValLeuGlyLeuAspAspIleHisArgAlaTrpArg
ThrPheValLeuArgValArgAlaGlnAspProProPro
GluLeuTyrPheValLysValAspValThrGlyAlaTyr
AspThrIleProGlnAspArgLeuThrGluValIleAla
SerIleIleLysProGlnAsnThrTyrCysValArgArg
TyrAlaValValGlnLysAlaAlaHisGlyHisValArg
LysAlaPheLysSerHisValLeuArgProValProGly
AspProAlaGlyLeuHisProLeuHisAlaAlaLeuGln
ProValLeuArgArgHisGlyGluGlnAlaValCysGly
AspSerAlaGlyArgAlaAlaProAlaPheGlyGly

FIG. 19

GCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGGCCACCCCCGCG ATG
1
met
10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC
20
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG
30
40
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GGC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT
50
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC
60
70
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC
80
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC
90
100
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC TTC GCG CTG CTG GAC GGG
110
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC
120

FIG. 20

MPRAPRCRAVRSLLRSHYREVLPLATFVRRLGPQGWRLVQRGDP
 AAFRALVAQCLVCVPWDARPPPAAPSFQVSCLELVARVLQRL
 CERGAKNVLAFGFALLDGAARGGPPEAFSTSVRSYLPNTVTDALR
 GSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLY
 QLGAATQARPPPHASGPRRRRLGCERAWNHSVREAGVPLGLPAPG
 ARRRGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRG
 PSDRGFCVVSAPPAEEATSLEGALSGTRHSHPSVGRQHAGPP
 STSRPPRPWDTPCPPVYAETKHFLYSSGDKEQLRPSFLLSSLRP
 SLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLEL
 LGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPQGSVAAPPE
 EDTDPRRLVQLLRQHSSPWQVYGFVRACLRLRLVPPGLWGSRHNE
 RRFLRNTKKFISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGC
 VPAAEHRLREEILAKFLHWLMSVYVVELLRSSFYVTETTFQKNR
 LFFYRKSVMWSKLSIGIRQHLKRVQLRELSEAEVRQHREARPAL
 LTRSRLRFIPKPDGLRPIVNM DYVVGARTFRREKRAERLTSRVKA
 LFSVLNYERARRPGLLGASVLGLDDIHRWRTFVLRVRAQDPPP
 ELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQ
 KAAHGHRVKAFAKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI
 EQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPOGSI
 LSTLLCSLCYGD MENKLFAGIRRDGLLLRLVDDFLLVTPHLLTHA
 KTFRLTLVRGVPEYGCVVNLKRTVVNFPVEDEALGGTAFVQMPA
 HGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR
 NMRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRF
 HACVLQLPFHQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL
 GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLGLSLRTAQ
 TQLSRKLPGTTLTALEAAANPALPSDFKTILD

FIG. 17

GGCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCTGCTCGAGCTGCTCAGGTCTTTCTT
 TTATGTCACGGAGACCACGTTTCAAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTG
 GAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT
 GTCGGAAGCAGAGGTGAGGCAGCATCGGGAAGCCAGGCCGCCCTGCTGACGTCCAGACT
 CCGCTTCATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCTGTTGG
 AGCCAGAAGCTTCCGCAGAGAAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCCT
 GTTCAGCGTGCTCAACTACGAGCGGGCGCGGCCGCCCTCCTGGGCGCCTCTGTGCT
 GGGCCTGGAGCATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCAGGA
 CCGGCCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC
 CCAGGACAGGCTCACGGAGGTATCGCCAGCATCATCAAACCCAGAACACGTACTGCGT
 GCGTCTGCTATGCCGTGGTCCAGAAAGCCGCCATGGGCACGTCCGCAAGGCCTTCAAGAG
 CCAGCTCCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCT
 GCAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGGCGGGGATTGCGGGGACGGG
 TGCTCCTGCGTTTGGTGGATGATTTCTTGTGTGGTGACACCTCACCTCACCCACGCGAAAA
 CCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACCTGCGGA
 AGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGA
 TGCCGGGCCACGGCCTATTTCCCTGGTGGCGCTGCTGCTGGATACCCGACCCCTGGAGG
 TGACAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACC
 GCGGCTTCAAGGCTGGGAGGAACATGCGTCCGAAACTCTTTGGGGTCTTGCGGCTGAAGT
 GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT
 ACAAGATCCTCCTGCTGCAGGCGTACAGGTTTACGCATGTGTGCTGCAGCTCCCATTTT
 ATCAGCAAGTTTGAAGAACCCACATTTTCTTCTGCGCGTCATCTCTGACACGGCCTCCC
 TCTGTACTCCATCTGAAAGCCAAGAAGCAGGGATGTGCTGGGGGCCAAGGGCGCCG
 CCGGCC7TCTGCCCTCCGAGGCCGTGCACTGGCTGTGCCACCAAGCATTTCTGCTCAAGC
 TGACTCGACACCGTGTACCTACGTGCCACTCTGGGGTCACTCAGGACAGCCAGACGC
 AGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGGCCGAGCCAAACCCG
 CACTGCCCTGACACTTCAAGACCATCTTGGACTGATGGCCACCCGCCACAGCCAGGCCG
 AGAGCAGACACCAGCAGCCCTGTACGCCCGGGCTCTACGTCCAGGGAGGGAGGGGCGGC
 CCACACCCAGGCCGTGACCCGCTGGGAGTCTGAGGCCGTGAGTGTGTTGGCCGAGGCCCT
 GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCGTGAGCGAGTGTCCAGCCAAGGGC
 TGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC
 AGGGCCAGCTTTTCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGGAATAGTCCATC
 CCCAGATTCCGCATTGTTTACCCCTCGCCCTGCCCTCCTTTGCTTCCACCCCCACCATC
 CAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGT
 GCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGGGTCCCTGTGGGTCAAATTTGGGG
 GGAGGTGCTGTGGGAGTAAAAATACTGAATATATGAGTTTTTTCAGTTTTTG0AAAAAAAAA
 AAAAAAAAAAAAAAAAAA

FIG. 18

06/9/2951

1 gcagcgctgc gtcctgctgc gcacgtggga agccctggcc cgggccaccc ccgcatgccc
 61 ggcgctcccc cgtgcccag cctgctgcgc agccactacc gcgaggtgct
 121 gccgctggcc acgttcgtgc ggcgctggg gcccagggc tggcggtgg
 181 ggaccggcg gctttccgc cgtgtgtgg ccagtgcctg gtgtgcgtgc
 241 acggccgccc cccgcccgc cctccttccg ccaggtgtcc tgcctgaagg
 301 ccgagtgctg cagaggctgt ggcgagcgg cgcgaagaac gtgctggcct
 361 gctgctggac gggggccgcg gggggccccc cgaggccttc accaccagcg
 421 cctgcccac acggtgaccg acgcactgcg ggggagcggg gcgtgggggc
 481 ccgctggggc gacgacgtgc tggttcacct gctggcacgc tgcgctctct
 541 ggctcccagc tgcgcctacc aggtgtgcgg gccgcgctg taccagctcg
 601 tcaggcccg ccccccgcac acgctagtgg accccgaagg cgtctgggat
 661 ctggaacct agcgtcaggg aggcgggggt cccctggggc ctggcagccc
 721 gagcgcggg ggcagtgcca gccgaagtct gccgttggcc aagaggccca
 781 tgcccttgag ccggagcggg cggccgttgg gcaggggtcc tggggccacc
 841 gcgtggaccg agtgaccgtg gtttctgtgt ggtgtcacct gccagacccc
 901 caccctcttg gaggtgctg tctctggcac gcccaactcc caccatccg
 961 gcaccacgc gggcccccac ccacatcgcg gccaccaagt cctggggaca
 1021 cccggtgtac gccgagacca agcacttcc ctactcctca ggcgacaagg
 1081 gccctccttc ctactcagct ctctgaggcc cagcctgact ggcgctcgga
 1141 gaccatcttt ctgggttcca ggcctggagt ccaggagact ccccgagct
 1201 gcccacgagc tactggcaaa tgcggccct gtttctggag ctgcttggga
 1261 gtgcccctac ggggtgctcc tcaagacgca ctgcccgtg cgagctgagg
 1321 agccggtgtc tgtgcccggg agaagcccca gggctctgtg gcggccccc
 1381 cacagacccc cgtgcctgg tgcagctgct ccgacgccc agcagccct
 1441 cggcttctg cgggcctgcc tgcgcccgt ggtgccccca ggcctctggg
 1501 caacgaacgc cgttctctca ggaacaccaa gaagtctatc tccctgggga
 1561 gctctcgtg caggagctga cgtggaagat gagcgtgagg gactgcgctt
 1621 gagcccaggg gttggctgtg ttccggccgc agagcacctg ctgctgagg
 1681 caagttccct cactggctga tgagtgtgta cgtcgtcag ctgctcaggt
 1741 tgtcacggag accacgtttc aaaagaacag gctctttttc taccggaaga
 1801 caagttgcaa agcattggaa tcagacagca cttgaagagg gtgcagctgc
 1861 ggaagcagag gtcaggcagc atcgggaagc caggcccgc ctgctgacgt
 1921 cttcatcccc aagcctgacg ggctgcggcc gattgtgaac atggactacg
 1981 cagaacgttc cgcagagaaa agagggccga gctctcacc tcgaggggta
 2041 cagcgtgctc aactacgagc gggcgccggc ccccgccctc ctggggcgct
 2101 cctggacgat atccacagg cctgctgctg cgtgtgagg gctgtgagg
 2161 gccgcctgag ctgtactttg tcaagggtga tgtgacgggc gcgtacgaca
 2221 ggacaggctc acggaggtca tcgccagcat catcaaacc cagaacacgt
 2281 tcggttatcc gtggtccaga agggcgccca tgggcacgtc cgcaaggcct
 2341 cgtctctacc ttgacagacc tccagcgta catcgacag ttcgtggctc
 2401 gaccagccc ctgagggatg ccgtcgtcat cgagcagagc tctcctctga
 2461 cagtggcctc ttgcagctct tctacgctt catgtgccac caccgctgct
 2521 caagtcctac gtccagtgc aggggatccc gtcagggtcc atcctctcca
 2581 cagcctgtgc tacggcgaca tggagaacaa gctgttgcg gggattcggc
 2641 gctcctgctg ttggtggatg atttcttgtt ggtgacacct caccctaccc
 2701 ctctctcagg accctggctc gaggtgtccc tgagtatggc tgcgtgggta
 2761 gacagtgggt aacttccctg tagaagacga gggcctgggt ggcacggctt
 2821 gccggccccc ggcctattcc cctggtgcgg cctgctgctg gataaccgga
 2881 gcagagcgac tactccagct atgccggac ctccatcaga gccagtctca
 2941 cggcttcaag gctgggagga acatgcgtcg caaactctt ggggtcttgc
 3001 tcacagcctg tttctggatt tgcaggtgaa cagcctccag acggtgtgca
 3061 caagatcctc ctgctgcagg cgtacaggtt tcacgcatgt gtgctgcagc
 3121 tcagcaagtt tggagaacc ccacattttt cctgcgcgtc atctctgaca
 3181 ctgctactcc atcctgaaag ccaagaacgc agggatgtcg ctgggggcca
 3241 cggccctctg cctcccgagg ccgtgcagt gctgtgccac caagcattcc
 3301 gactcgacac cgtgtcacct acgtgccact cctgggggta ctgaggacag
 3361 gctgagtcgg aagctcccgg ggacgacgt gactgccctg gaggccgag
 3421 actgcccctc gacttcaaga ccatcctgga ctgatggcca cccgcccaca
 3481 gagcagacac cagcagccct gtcagccgg gctctacgtc ccagggaggg
 3541 cacaccagg cccgaccgc tgggagctg aggcctgagt gagtgtttgg
 3601 catgtccggc tgaaggctga gttgtccggt gaggcctgag cgagtgtcca
 3661 gagtgtccag cacacctgcc gtcttcaact cccacaggc tggcgctcgg
 3721 gggccagctt ttccctacca ggagcccggc ttccactccc cacataggaa
 3781 ccagatctgc cattgttcc cctcgcctt gccttccacc gcttccacc
 3841 aggtggagac cctgagaagg accctgggag cctctgggaat ttggagtgc
 3901 cctgtacac aggcgaggac cctgcacct gatgggggtc cctgtgggtc
 3961 gaggtgctgt gggagtaaaa tactgaatat atgagttttt cagttttgaa

FIG. 16

156216/20

4666 gctgacccccaaagcagcactataggtttcttagtaaaagtaaaaaataatctctcgttatttagtttttgattgacttggtct 4745
4746 ttatccttataacttttaagaaagattgacagtggttgctgactactgcccacatgcccattaaacgggagtggttaaca 4825
4826 ttaaaagtaatacatgaggttaattctcctttcatttagaataaggaagtggttttctataatgaataatgcccgcacta 4905
4906 atgcaaaaagacgaagattatcttctaacaaggggattcaagcatatccgaaggaaaaagagagtaatataccagtggtt 4985
4986 gttgaagaaagcaggataaatttggaacaagcttctgcagatgacaggctaaaattttggtgaccgaatttttggtaaaagc 5065
5066 cccaggttatccatggtgcccggccttgctactgagacgaaaaaactaaaggatagtttgaataactaataagctcattta 5145
5146 atgtcttataaaggttttttttctcgtgacttcaatttttgcatgggtgaaaaaagaaatagtgtaagccattattggat 5225
5226 tccgaaatagccaaattttcttggttccctcaagcgggaagtctaaagaaacttattgaagcttatgaggttcaaaaactcc 5305
5306 tccgtatttaaaaggaggaattctccaccgatgaggaaatggatagcttatcagctgctgaggagaagcctaatttttgc 5385
5386 aaaaaaagaaatatcatgtgggagacatctcttgatgaatcagatgcggagagtatctccagcggatccttgatgtcaata 5465
5466 acttctatttctgaaatgtatgggtccctactgtctcgcttcgacttctcgtagctctacgcaggttaagtgaaccaaagggtacc 5544

FIG. 15
(CONTINUED)

156716/80

18/34

3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT 3960
799 R M P F F G F S V N M R S L D T L L A C 818

3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG 4020
819 P K I D E A L F N S T S V E L T K H M G 838

4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtaactgaataatagctgacaaataatcag A TCG 4089
839 K S F F Y K I L R S 848

4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT 4149
849 S L A S F A Q V F I D I T H N S K F N S 868

4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA 4209
869 C C N I Y R L G Y S M C M R A Q A Y L K 888

4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagtacttattttaactaga 4274
889 R M K D I F I P Q R M F I T D 903

4275 aaagtcattaattaaccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC 4339
904 L L N V I G R K I W K K L A 917

4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401
918 E I L G Y T S R R F L S S A E V K W 935

4402 ggtctcgagacttcagcaaatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA 4468
936 L F C L G M R D G L K 946

4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT 4528
947 P S F K Y H P C F E Q L I Y Q F Q S L T 966

4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA ATA 4588
967 D L I K P L R P V L R Q V L F L H R R I 986

4589 GCT GAT TAA tgtcattttcaattattatatatacatcctttattactggtgtctttaaacaaatattattactaagtata 4665
987 A D * 989

FIG. 15
(CONTINUED)

156216/80

3089 tatataatgcgcgattccctcattattaattttgcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 3155
582 R K K Y F V R I D I 591

3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 3215
592 K S C Y D R I K Q D L M F R I V K K L 611

3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3275
612 K D P E F V I R K Y A T I H A T S D R A 631

3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagtttttttttcattggaattttttaacaa 3343
632 T K N F V S E A F S Y F 643

3344 attcttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3405
644 D M V P F E K V V Q L L S M K T 659

3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3465
660 S D T L F V D F V D Y W T K S S E I F 679

3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgtgaattgtaataaca 3532
680 K M L K E H L S G H I V K 692

3533 ctaatgaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA 3593
693 I G N S Q Y L Q K V G I P Q G S 708

3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3653
709 I L S S F L C H F Y M E D L I D E Y L S 728

3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3713
729 F T K K K G S V L L R V V D D F L F I T 748

3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgtgtcattcc 3777
749 V N K K D A K K F L N L S L R G 764

3778 taagttctaaccgttggaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3840
765 F E K H N F S T S L E K T V 778

3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAA 3900
779 I N F E N S N G I I N N T F F N E S K K 798

FIG. 15

(CONTINUED)

156216/80

16/34

2268 TTT GAG ATA ATA TTA AAA G gattgtataaaatttattaccactaacgattttaccag AC CTC GAA ACT 2336
 366 F E I I L K D L E T 375

2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
 376 F L K L S R Y E S F S L H Y L M S N I K 395

2397 gtaatatgccaaattttttaccatttaataacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465
 396 I S E I E W L V L G 405

2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
 406 K R S N A K M C L S D F E K R K Q I F A 425

2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
 426 E F I Y W L Y N S F I I P I L Q S F F Y 445

2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
 446 I T E S S D L R N R T V Y F R K D I W K 465

2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
 466 L L C R P F I T S M K M E A F E K I N E 485

2706 gtttttaaaagtatttttttgcacaaagctaataattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
 486 N N V R M D T Q K T 495

2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835
 496 T L P P A V I R L L P K N T F R L I T 515

2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtatttaatttttgggtcatcaatgtacttttacttctaatttatta 2906
 516 N L R K R F L I K 524

2907 ttagcag ATG GGT TCA AAC AAA AAT ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
 525 M G S N K K M L V S T N Q T L R P V 542

2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
 543 A S I L K H L I N E S S G I P F N L E 562

3028 GTT TAC ATG AAG CTT ACT CTT AAG AAG GAT CTT AAG CAC CGA ATG TTT GG gtaat 3088
 563 V Y M K L L T F K K D L L K H R M F G 581

FIG. 15
(CONTINUED)

156716/20

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA 1529
 129 D L V S T F P N Y L I S I L E S K N W Q 148
 1530 CTT TTG TTA GAA AT gtaaataccgggttaagatgttgcgcaactttgaacaagactgacaagtatat T ATC GGC 1601
 149 L L L E I I G 155
 1602 AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661
 156 S D A M H Y L L S K G S I F E A L P N D 175
 1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721
 176 N Y L Q I S G I P L F K N N V F E E T V 195
 1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781
 196 S K K R K R T I E T S I T Q N K S A R K 215
 1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841
 216 E V S W N S I S I S R F S I F Y R S S Y 235
 1842 AAG AAG TTT AAG CAA G gtaactaatactgttattcccttcataactaatttttag AT CTA TAT TTT AAC 1907
 236 K K F K Q D L Y F N 245
 1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967
 246 L H S I C D R N T V H M W L Q W I F P R 265
 1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027
 266 Q F G L I N A F Q V K Q L H K V I P L V 285
 2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087
 286 S Q S T V V P K R L L K V Y P L I E Q T 305
 2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147
 306 A K R L H R I S L S K V Y N H Y C P Y I 325
 2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207
 326 D T H D D E K I L S Y S L K P N Q V F A 345
 2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267
 346 F L R S I L V R V F P K L I W G N Q R I 365

FIG. 15
(CONTINUED)

156216/20

1 ggtaccgattacttcttcttccataagctaattggtcttccctcgaaagctcctaataatctctggaaaatattttacaaga 80
 81 actcaataacaatacccaagtcacaaattccaatatgaaggtgtatttagtgatcgataatatttctatttctcggtcgta 160
 161 ccaagtataaggacaaaagaacaacttcttccccctaaagacttttacttttataatttacttttcaaatattttcg 240
 241 ggttcgcttacttttaactgtgtgactgttttagtctacttctagccaacgcgtgtttctaccccgctcatggatat 320
 321 agctcttgagtagctcacagaaatccctacaaatcttctgatgagactatattagatcaattacagtcggtgatattc 400
 401 ttaacatggagcccttacacttttagatgagtcacgtcgcatgaggtattttggatcatccaacgtttgcccctgaaaa 480
 481 gttgataattatttgcaaatcatgtccttagtggtggtgtaacccgcgaagtttttctgattgcttgcacacgtctagcatg 560
 561 attgagatattcaaaaattctatccactacaactcctttaaagcggttttatttcttatttcttctcattctgtt 640
 641 ccaaatatgtatcattatgtattaggtttttccggttttactcctggaatcgtaacttttccactattcccccaatga 720
 721 ataacttaaatagtttcgcttataattgatagtagtagaagaattggtgattctactcgtgtaattgtattagtttaaa 800
 801 gatactttgcaaaacattttagctatcattatataaaaaaacctataattataaataattcaatatttgcggtc 880
 881 actatttatttaaacgttatgatcagtaggacatttgcatatatatagttatgcttaatggttacttgaactgc 958

 959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
 1 M T E H H T P K S R I L R F L E N Q Y V 20

 1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
 21 Y L C T L N D Y V Q L V L R G S P A S S 40

 1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
 41 Y S N I C E R L R S D V Q T S F S I F L 60

 1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
 61 H S T V V G F D S K P D E G V Q F S S P 80

 1199 AAA TGC TCA CAG TCA GAG gtatatatattttgttttggattttttctattcgggatagctaataatgggcag 1272
 81 K C S Q S E 86

 1273 CTA ATA GCG AAT GTT GTA AAA CAG ATG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA 1332
 87 L I A N V V K Q M F D E S F E R R N L 106

 1333 CTG ATG AAA GGG TTT TCC ATG gtaagggtatttctaattgtgaaatatttaccctgcaattactgtttcagaaga 1405
 107 L M K G F S M 113

 1406 ttgtatttaaccgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT 1469
 114 N H E D F R A M H V N G V Q N 128

FIG. 15

2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
 2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
 2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
 2551 GACTAGTTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
 2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
 2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
 2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
 2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
 2801 CATTATTTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAACTCT
 2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
 2901 AATACAAGGA CCACTTTAAG AAGAAGCTAG CTATGAGCAG TATGATCGAC
 2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
 3001 CCTTGTGTGC AATATTAAGG ATACAATTTT TGGAGAGGAG CATTATCCAG
 3051 ACTTTTTTCT TAGCACACTG AAGCACTTTA TTGAAATATT CAGCACAAAA
 3101 AAGTACATTT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGGCAA
 3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
 3201 CTATTCTAAC TTATTTTGGG AAGTTAATTT TCAATTTTTG TCTTATATAC
 3251 TGGGGTTTTG GGGTTTTGGG GTTTTGGGG

FIG. 13
(CONTINUED)

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQKVIRC RNQSQSHYKD
 51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
 101 SSSDVSDRQK LQCFGFQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
 151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGGA
 201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNMMKSRTTR IFYCTHFNRN
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAM
 301 LEKVKDFNFN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHელი
 401 KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKLK RWIFEDLVVS
 451 LIRCFYFVTE QQKSYSKTTY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV
 501 EEWKKS LGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTTNTKLL
 551 NSHMLMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN
 651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ
 701 RNYFKKDNLL QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFY
 751 ATLEESSLGF LRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIKEL
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
 851 SIDMKTALM PNINLRIEGL CTLNLMNMQT KKASMWLKKK LKSFLMNNIT
 901 HYFRKTITTE DFANKTLNKL FISGGYKYM CAKEYKDHFK KNLAMSSMID
 951 LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
 1001 KYIFNRVCM I LKAKEAKLKS DQCQSLIQYD A

FIG. 14

08/9/2951

08/9/295/

1	AAAACCCCAA	AACCCCAAAA	CCCCTTTTAG	AGCCCTGCAG	TTGGAAATAT
51	AACCTCAGTA	TTAATAAGCT	CAGATTTTAA	ATATTAATTA	CAAAACCTAA
101	ATGGAGGTTG	ATGTTGATAA	TCAAGCTGAT	AATCATGGCA	TTCACTCAGC
151	TCTTAAGACT	TGTGAAGAAA	TTAAAGAAGC	TAAAACGTTG	TACTCTTGGA
201	TCCAGAAAGT	TATTAGATGA	AGAAATCAAT	CTCAAAGTCA	TTATAAAGAT
251	TTAGAAGATA	TTAAAAATATT	TGCGCAGACA	AATATTGTTG	CTACTCCACG
301	AGACTATAAT	GAAGAAGATT	TTAAAGTTAT	TGCAAGAAAA	GAAGTATTTT
351	CAACTGGACT	AATGATCGAA	CTTATTGACA	AATGCTTAGT	TGAACCTCTT
401	TCATCAAGCG	ATGTTTCAGA	TAGACAAAAA	CTTCAATGAT	TTGGATTTC
451	ACTTAAGGGA	AATCAATTAG	CAAAGACCCA	TTTATTAACA	GCTCTTTCAA
501	CTCAAAGCA	GTATTTCTTT	CAAGACGAAT	GGAACCAAGT	TAGAGCAATG
551	ATTGGAAATG	AGCTCTTCCG	ACATCTCTAC	ACTAAATATT	TAATATTCCA
601	GCGAACTTCT	GAAGGAACTC	TTGTTCAATT	TTGCGGGAAT	AACGTTTTTG
651	ATCATTTGAA	AGTCAACGAT	AAGTTTGACA	AAAAGCAAAA	AGGTGGAGCA
701	GCAGACATGA	ATGAACCTCG	ATGTTGATCA	ACCTGCAAAT	ACAATGTCAA
751	GAATGAGAAA	GATCACTTTC	TCAACAACAT	CAACGTGCCG	AATTTGGAATA
801	ATATGAAATC	AAGAACCAGA	ATATTTTATT	GCACTCATTT	TAATAGAAAT
851	AACCAATTCT	TCAAAAAGCA	TGAGTTTGTG	AGTAACAAAA	ACAATATTTT
901	AGCGATGGAC	AGAGCTCAGA	CGATATTCAC	GAATATATTC	AGATTTAATA
951	GAATTAGAAA	GAAGCTAAAA	GATAAGGTTA	TCGAAAAAAT	TGCCTACATG
1001	CTTGAGAAAG	TCAAAGATTT	TAACCTCAAC	TACTATTTAA	CAAAATCTTG
1051	TCCTCTTCCA	GAAAATTGGC	GGGAACGGAA	ACAAAAAATC	GAAAACCTGA
1101	TAAATAAAAC	TAGAGAAGAA	AAGTCGAAGT	ACTATGAAGA	CTGTGTTAGC
1151	TACACAACAT	ATAATAAATG	CGTCACACAA	TTTATTAATG	AATTTTCTTA
1201	CAATATACTC	CCCAAAGACT	TTTTGACTGG	AAGAAACCGT	AAGAATTTTC
1251	AAAAGAAAGT	TAAGAAATAT	GTGGAACATA	ACAAGCATGA	ACTCATTCAC
1301	AAAAACTTAT	TGCTTGAGAA	GATCAATACA	AGAGAAATAT	CATGGATGCA
1351	GGTTGAGACC	TCTGCAAAGC	ATTTTTATTA	TTTTTGATCAC	GAAAACATCT
1401	ACGTCTTATG	GAAATTGCTC	CGATGGATAT	TCGAGGATCT	CGTCGTCTCG
1451	CTGATTAGAT	GATTTTCTTA	TGTCACCGAG	CAACAGAAAA	GTTACTCCAA
1501	AACCTATTAC	TACAGAAAGA	ATATTGGGA	CGTCATTATG	AAAATGTCAA
1551	TCGCAGACTT	AAAGAAGGAA	ACGCTTGCTG	AGGTCCAAGA	AAAAGAGGTT
1601	GAAGAATGGA	AAAAGTCGCT	TGGATTTGCA	CCTGGAAAAC	TCAGACTAAT
1651	ACCGAAGAAA	ACTACTTTCC	GTCCAATTAT	GACTTTCAAT	AAGAAGATTG
1701	TAAATTCAGA	CCGGAAGACT	ACAAAATTAA	CTACAAATAC	GAAGTTATTG
1751	AACTCTCACT	TAATGCTTAA	GACATTGAAG	AATAGAATGT	TTAAAGATCC
1801	TTTTGGATTG	GCTGTTTTTA	ACTATGATGA	TGTAATGAAA	AAGTATGAGG
1851	AGTTTGTTTG	CAAATGGAAG	CAAGTTGGAC	AACCAAAACT	CTTCTTTGCA
1901	ACTATGGATA	TCGAAAAGTG	ATATGATAGT	GTAAACAGAG	AAAAACTATC
1951	AACATTCCTA	AAAACACTA	AATTACTTTC	TTTCAATTTT	TGGATTATGA
2001	CTGCACAAAT	TCTAAAGAGA	AAGAATAACA	TAGTTATCGA	TTTCAAAAAC
2051	TTTAGAAAGA	AAGAAATGAA	AGATTATTTT	AGACAGAAAT	TCCAGAAGAT
2101	TGCACTTGAA	GGAGGACAAT	ATCCAACCTT	ATTCAGTGTT	CTTGAAAATG
2151	AACAAAATGA	CTTAAATGCA	AAGAAAACAT	TAATTGTTGA	AGCAAAGCAA
2201	AGAAATTATT	TTAAGAAAGA	TAACCTACTT	CAACCAGTCA	TTAATATTTG
2251	CCAATATAAT	TACATTAAGT	TTAATGGGAA	GTTTTATAAA	CAAACAAAAG
2301	GAATTCTTCA	AGGTCTTTGA	GTTTCATCAA	TTTTGTCTATC	ATTTTATTAT
2351	GCAACATTAG	AGGAAAGCTC	CTTAGGATTC	CTTAGAGATG	AATCAATGAA

FIG. 13

181 GGACCCGGCGGCTTTCCGCGCGCTGGTGGCCAGTGCCTGGTGTGCGTGCCCTGGGACGC
CCTGGGCGCGCGAAAGGCGCGCGACCACCGGGTCACGGACCACACGCACGGGACCCTGCG

NFkB_CS1
GGGRQTYYYQC
NFkB-MHC-I.2
TGGGCTTCCCC

241 ACGGCCGCCCCCGCCGCCCCCTCCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCG
TGCCGGCGGGGGGCGCGGGGAGGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGCAGGC

Intron1

301 GCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTC
CGACCCCAACTCCCGCCGCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

NFkB_CS1
GGGRQTYYYQC
NFkB_CS2
RGGGRMTYYCC
Topo_II_cleavage_site
RNYNNCNGYNGKTNINY
*****>

361 AGGGCGCTTCCCCCGCAGGTGTCTTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGG
TCCCGCGAAGGGGGCGTCCACAGGACGGACTTCCTCGACCACCGGGCTCACGACGTCTCC

FIG. 12

08/9/2951

156216/80

10/34

Telomerase Specific Motifs

	MOTIF T	MOTIF T'
TRT con	Wl	
hTRT	546 WLMSVVVVELLSFFVVTETTFQKNRLFFYRKSVWSKLQSIGI 13 EAEVR	E V
spTRT	429 WLYNSFIIPILQSFYITESDLNRNRTVYFRKDIWKLLCRPFI 12 ENNVR	
Ea_p123	441 WIFEDLVVSLIRCFYVTEQQKSYKTYVYRKNIWDVIMKMSI 12 EKEVE	
Sc_Est2	366 WLFRLIPKIIQTFYCYCTEISSTVT.IVYFRHDTWNKLITPFI 9 ENNVC	

Telomerase RT Motifs (Fingers)

	MOTIF 1	MOTIF 2	MOTIF A	MOTIF B'
TRT con	R ipkk	fr i	p lyf	Y q
hTRT	11 SRLRFIPKPDG 0 LRPIV	69 PELYFVKVDVTGAYDTI 104 YVQCQIPQGSILSTLLCSLCY	D cyd i	Y q
spTRT	10 AVIRLLPKKNT 0 FRLIT	66 RKKYFVRIDIKSCYDRI 99 YLQKVGIPQGSILSSFLCHFYM		
Ea_p123	10 GKRLIPKKT 0 FRPIM	67 PKLFFATMDIEKCYDSV 117 YKQTKGIPQGLCVSSILSSFY		
Sc_Est2	13 SKMRIIPKKS 2 FRIIA	68 PELYFMKFDVKSCYDSI 85 YIREDGLFQGSLSAPIVDLVY		
RT con	p hh h k	hR h	h hDh AF h	hpQG pp hh h
			GY	

Telomerase RT Motifs (Palm, Primer Grip)

	MOTIF C	MOTIF D	MOTIF E
TRT con	l1lrl	DDfL it	g
hTRT	15 LLLRLVDDFLVT 15	GVPEYGCVVNLRKTVV 24	WCGLLLDTRTL 192
spTRT	16 VLLRVVDDFLFIT 15	GFEKHNFSTSLKTVI 22	FFGFSVNMRS 176
Ea_p123	24 LLMRLTDDYLLIT 15	VSRENGFKFNMKKLQT 28	WIGISIDMKTL 174
Sc_Est2	18 LILKLADDFLIIS 15	GFQKYNKANRDKILA 25	WKHSSTMMNFH 141
RT con	h y Ddh	h h ck h	hLG h
	F		

FIG. 11

9/34

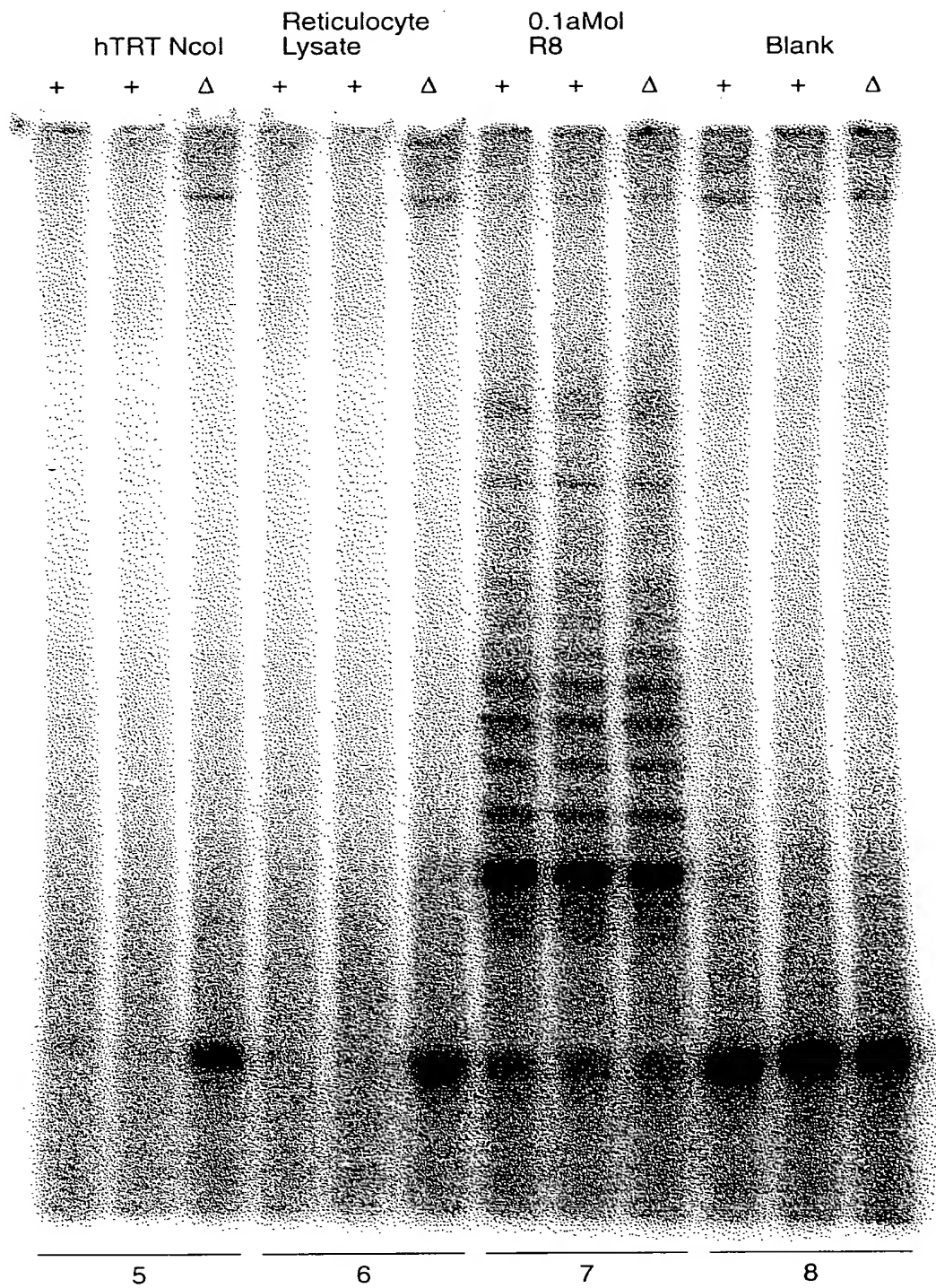


FIG. 10B

08/9/2957

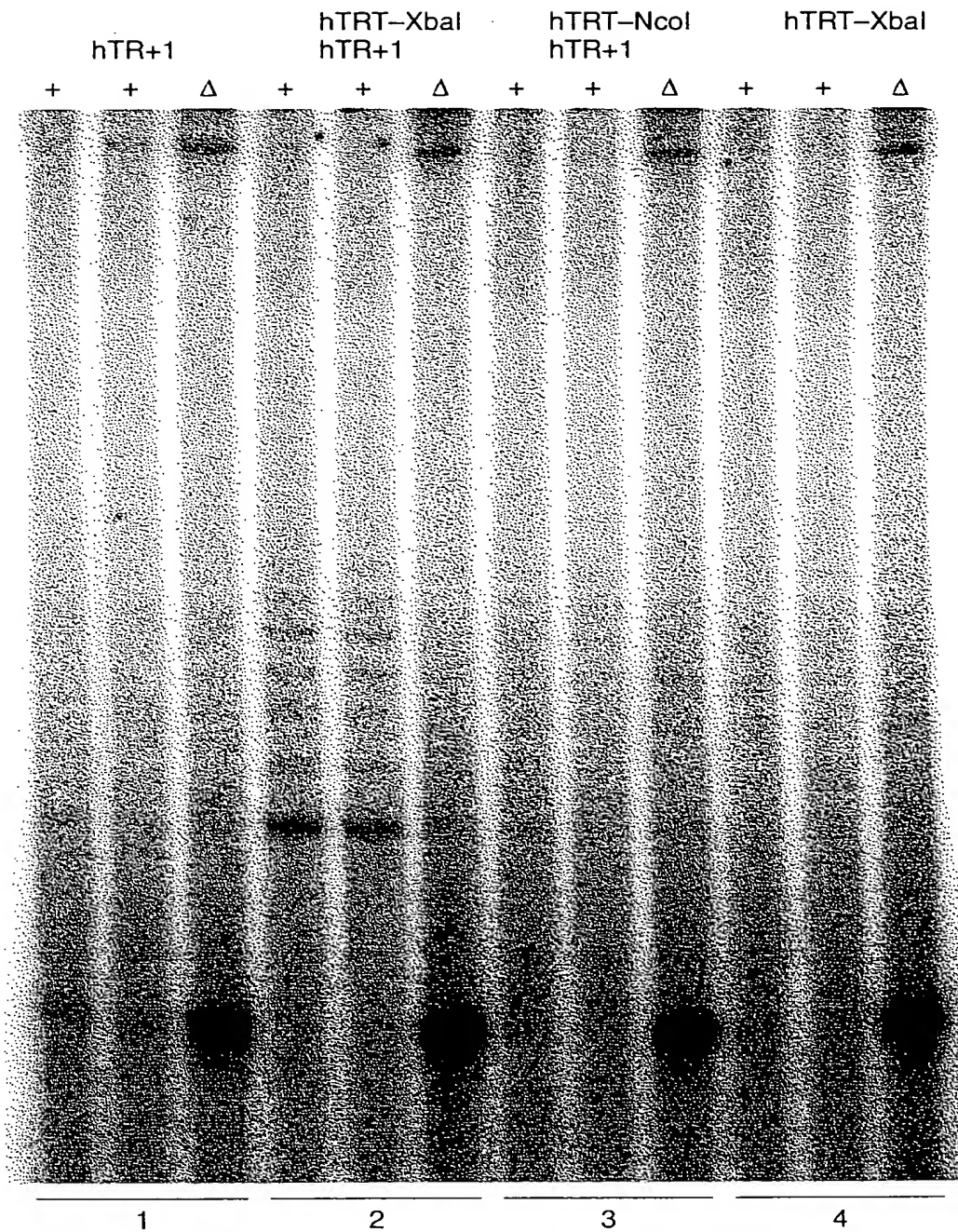


FIG. 10A

08/9/2951

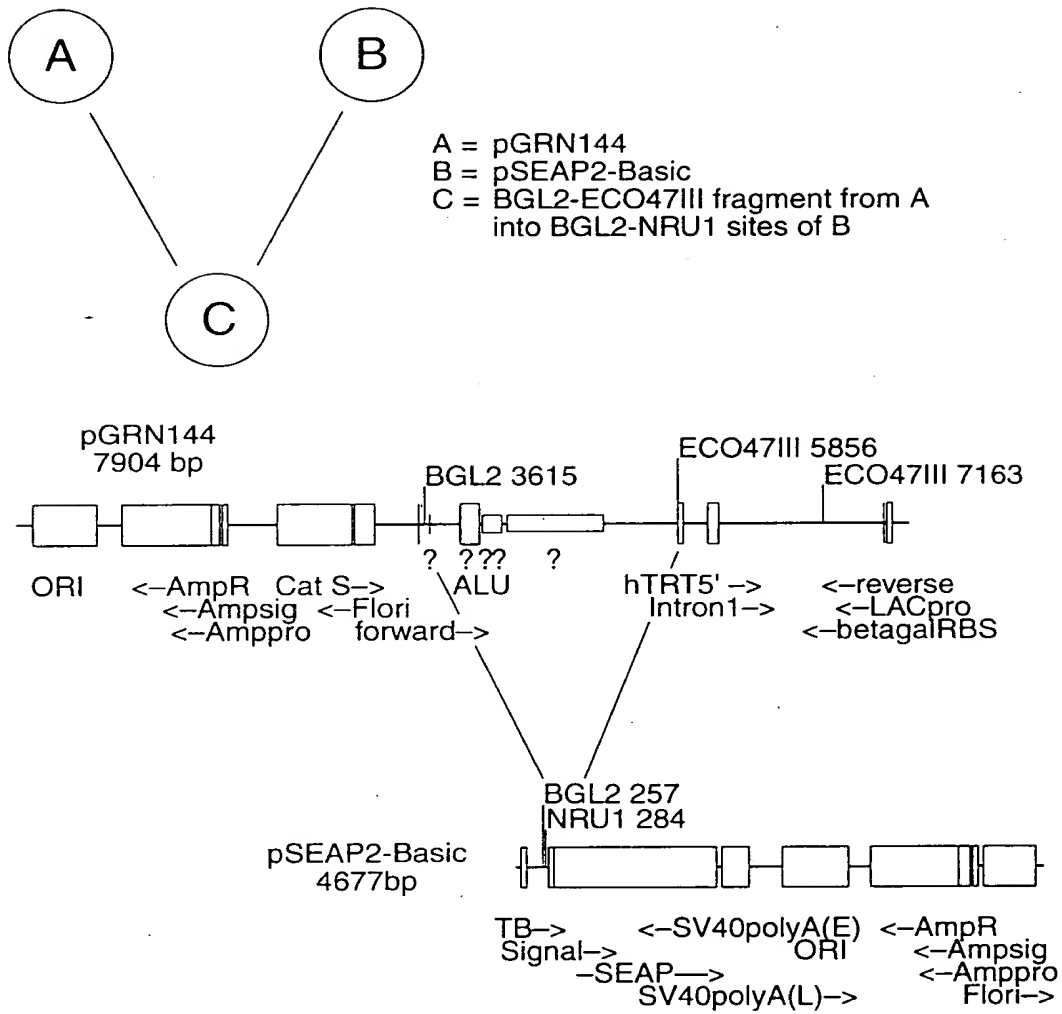


FIG. 9

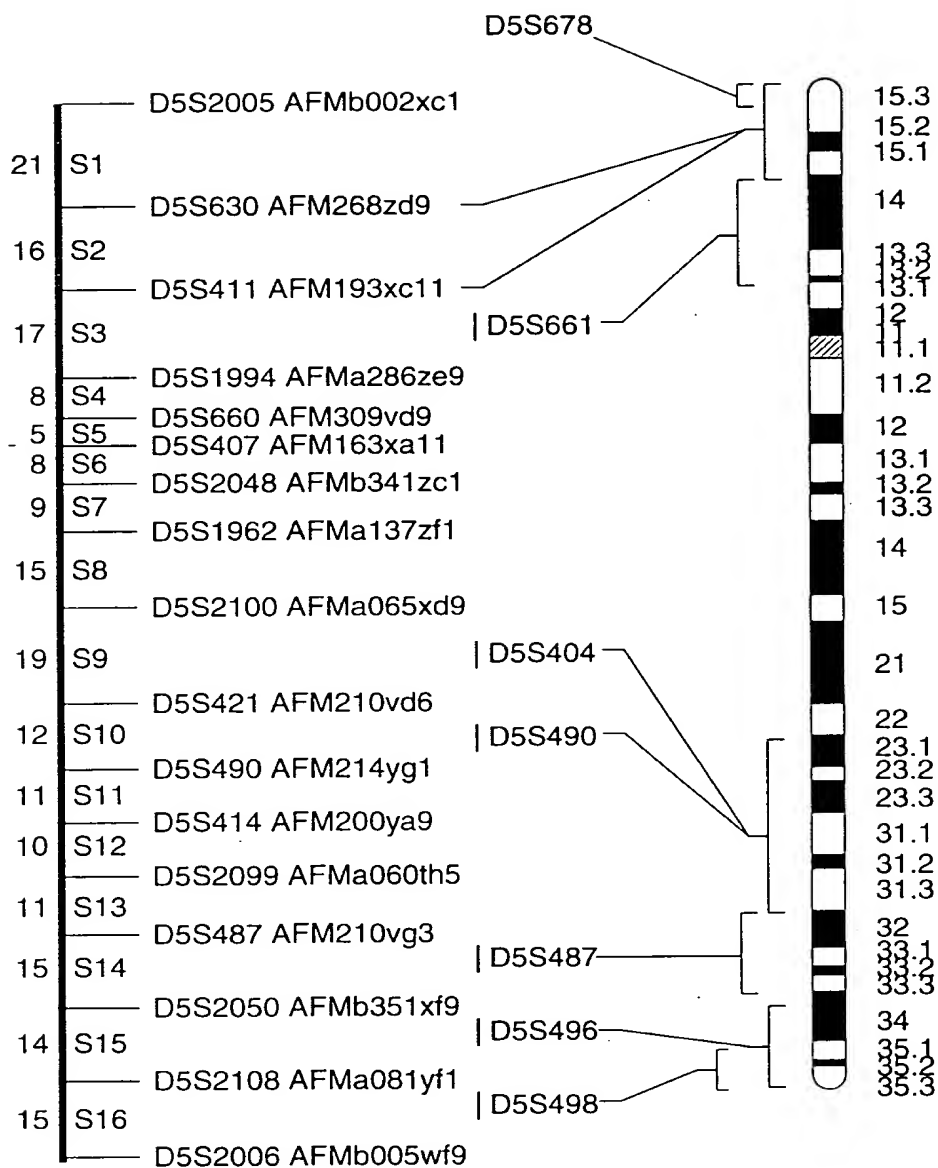


FIG. 8

5/34

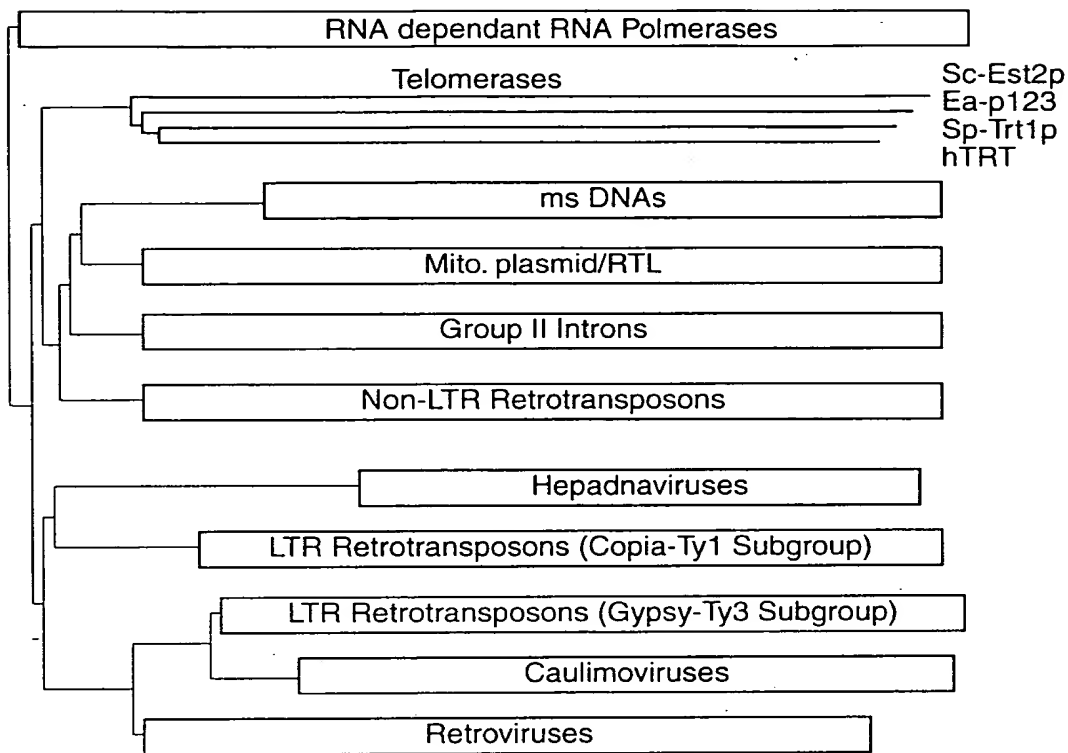


FIG. 6

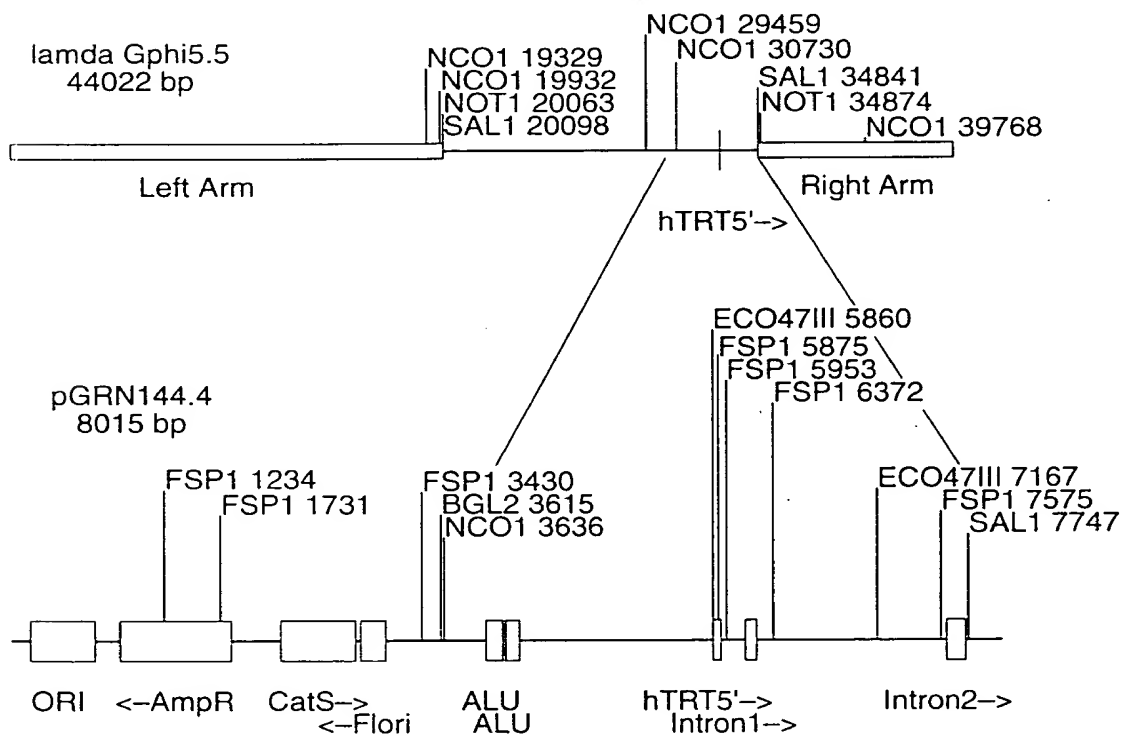


FIG. 7

08/9/2951

156216/80

		Motif T									
TRT con		WL	hh	hh	h	h	h	h	h	h	h
Sp_Trt1p		429	WLYNSFIPILOQ	FFFY	TE	p	p	Y	RK	W	L
hTRT		546	WLMVSVVVELLR	SSFF	VTET	TFQKNR	FFYRKS	VSWSKL	QSIGIR	HLK	10
Ea_p123		441	WIFEDLVSLIR	CCFF	YVTE	QQKSYS	KTYVRK	NIWDVIMK	MSIADL	KK	8
Sc_Est2p		366	WLFRLQILPKII	QTFY	CTEIS	STVT-IV	YFRHDT	WNKLIT	PPFIVE	YFK	8
		Motif A									
TRT con			h	h	h	h	h	h	h	h	h
Sp_Trt1p		NNVRMDTQ	KTLLP	PAVIR	LLPK	NT-0	FR	LTNL	RKRFL	IKMG	SNKMLVSTNQTL
hTRT		EVQRHREAR	PALLT	SRLR	FIPK	PDG-0	LR	PIVND	YVVGART	FRREK	RAERLTSRV
Ea_p123		KEVEEWKKS	GLGFA	PGKLR	LIPK	KT-0	FR	PIVND	YVVGART	FRREK	RAERLTSRV
Sc_Est2p		CRNHNSY	TLSNF	NHSMR	RIIP	KSNN	1	FR	ITAI	PCGA	DEEEFTIYKENHKNAIQP
		Motif B									
TRT con			p	hh	h	K	h	R	h	h	h
Sc_a1		LSNELGTG	KFKPM	RIVN	IPK	PGG	0	IR	PLSV	GNPR	DKIVQEV
Dm_TART		SILRIGY	YPDA	WKHA	QVMIL	KPGKS	6	YR	PI	SL	SGLSKMFERLL
HIV-1		EGKISKI	GPEN	PYNT	VPFA	IKK	DST	1	WR	KLVD	FRELNKRTQDFWEVQLGIPHPAG
		Motif C									
TRT con		K	Y	Q	GIP	QGS	LS	hL	h	Y	DL
Sp_Trt1p		SOYLQKV	GIP	QGS	ILS	SLFCH	FYMED	LID	EYLS	FT	6
hTRT		KSYVQC	QGIP	QGS	ILS	TLCS	LCYGD	ME	NKLF	AGI	5
Ea_p123		KFYKQTK	GIP	QGL	CVSS	ILS	SSFY	ATL	EES	SLGFL	14
Sc_Est2p		KCYIRE	DGLF	QGS	SL	SAPI	VDLV	VD	DL	LEFYSEFK	8
		Motif D									
TRT con			h	h	h	h	h	h	h	h	h
Sp_Trt1p		TYHKP	MLGL	PQGS	LIS	PILC	NI	VM	TLVD	NW	LEDYI
hTRT		RAGQIG	AGVP	QGS	NLGP	ILYIF	SSDM	PLPH	IYHP	7	LSTYAD
Ea_p123		GIRYQ	YNVLP	QGS	KGSP	ATFQ	SSMT	KILE	PF	KKQN	4
Sc_Est2p		hPQG	pp	hh	h	h	h	h	h	h	h
		Motif E									
TRT con			h	h	h	h	h	h	h	h	h
Sp_Trt1p		KKRMP	FFG	FSV	181	KKRMP	FFG	FSV	181	KKRMP	FFG
hTRT		HGLF	PCW	GLLL	197	HGLF	PCW	GLLL	197	HGLF	PCW
Ea_p123		QDYCD	WIG	ISI	179	QDYCD	WIG	ISI	179	QDYCD	WIG
Sc_Est2p		KELEV	WKHS	ST	146	KELEV	WKHS	ST	146	KELEV	WKHS

FIG. 4

3/34

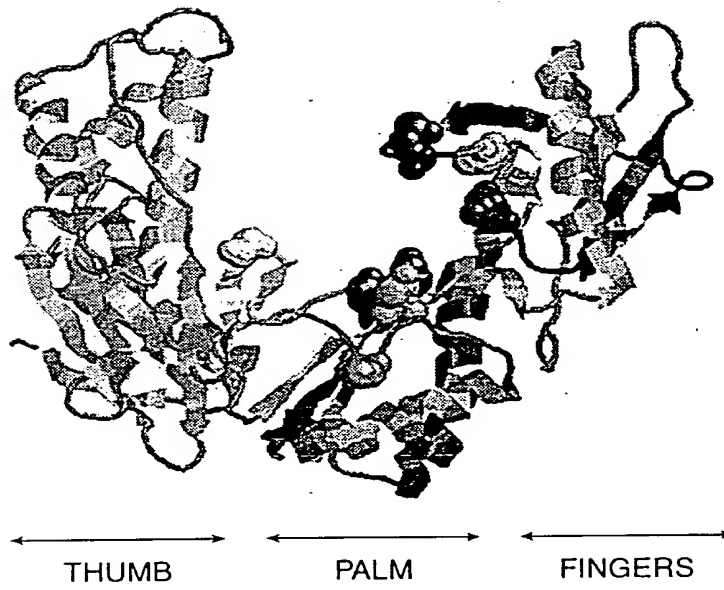
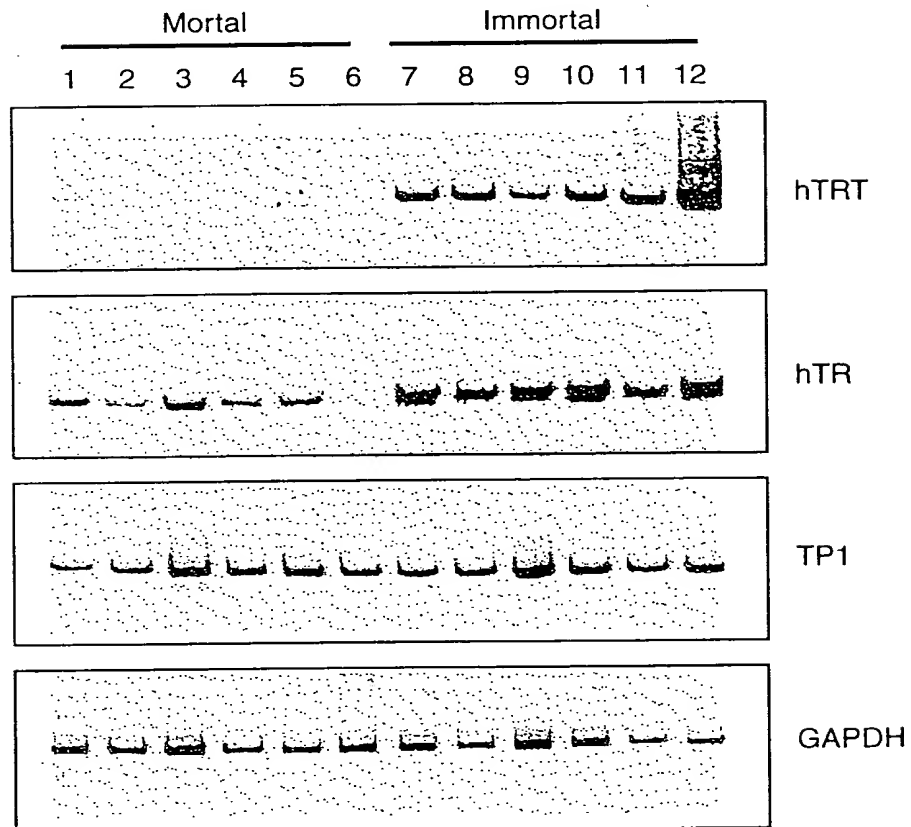


FIG. 3



156216/80

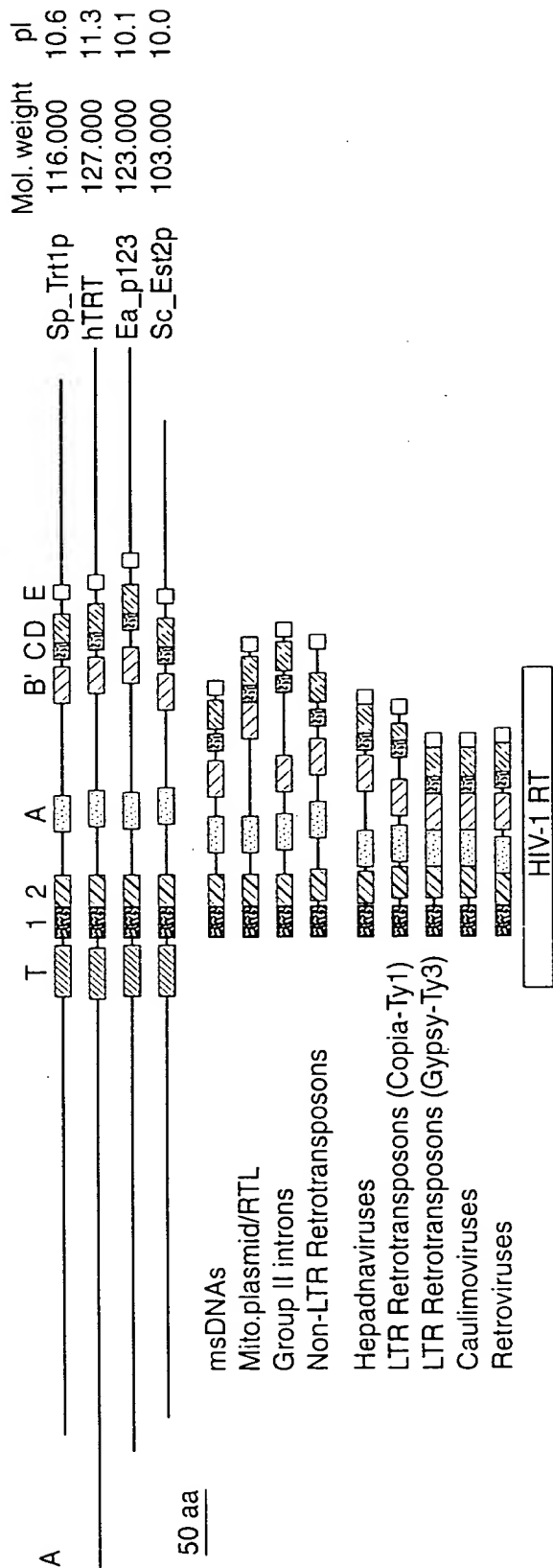


FIG. 2

human
tezi
EST2
p123

AKFLHWMVSVVVELLSRFFVYVTTFTQKNR
ISEIWLVLGKRSNAKMC LSDFEKRQIFAEFIWLYNSFIPILOQFFYITESSDLNR
LKDFRWLFISD--IWFTHFNENLQLAICFISWLFRLQIPKIIQTFYCYTESSTVT-
TREISWNQVET-SAKHFYYFDHEN-IYVLWKLRRWIFEDLVVSLIRCFWFVTEQQKSYSK

Motif 0
**** *

[illegible]

Motif 2

human	RPVNMDEVVGARTFRREKRAERLTSRVKALF-SVLNYERA	
tez1	RLITN-LRKRFLIKMSNKKMLVSTNQTLRPVASILKHLINESSGIPFNLEVYMKLLTF	
EST2	RIIAIPCGADEEFTIYKENHKNAIQPTQKILEYLRNKRPTSTFKIYSPQIADRIKEF	
p123	RPITFNKKIVNSDRKTTKLTNTKLLNSHLMKTLKN-RMFKDPFGFAVFNYYDDVMKKY	* * *

tezl KDLLKHRMFR-KKYFVRIDIKSCYDRIKQDLMFRIKK-KLKDPFVIRKYATIHATS
EST2 KQRLKKFNVLPELYFMKFDVKSCYSIPRMECMRILKD-ALKNENGFFVRSQYFFNTN
p123 EEFVCKWKQVGQPKLFFATMDIEKCYDSVNREKLSTFLKTKLLSSDFWIMTAQILKRKN

FIG. 1